FACULTEIT LETTEREN
EN WIJSBEGEERTE

The role of expressivity and productivity in (re)shaping the constructional network

Emmeline Gyselinck

Proefschrift voorgelegd tot het behalen van de graad van Doctor in de Taalkunde

UNIVERSITEIT
GENT

Promotor	Prof. dr. Timothy Colleman Vakgroep Taalkunde
Decaan	Prof. dr. Marc Boone
Rector	Prof. dr. Rik Van de Walle

FACULTEIT LETTEREN
EN WIJSBEGEERTE

The role of expressivity and productivity in (re)shaping the constructional network

A corpus-based study into synchronic and diachronic variation in the intensifying fake reflexive resultative construction in 19th to 21st Century Dutch

Emmeline Gyselinck

Proefschrift voorgelegd tot het behalen van de graad Doctor in de Taalkunde
2018

UNIVERSITEIT
GENT

Acknowledgements

Beste lezer

Hier is het dan, mijn doctoraat. Wie had dat gedacht. Ik zie mezelf als BA3-studente nog zitten in het kantoor van Timothy Colleman (voor mij toen nog "professor Colleman") met de vraag: 'Ik zou graag doctoreren, hoe doe je dat eigenlijk?' Er volgde een uitgebreide uitleg over de verschillende manieren waarop je kon doctoreren en, vooral, de vereisten waaraan je moest voldoen om überhaupt in aanmerking te komen. Als taalkundige *rookie* moest ik me op dat moment natuurlijk nog bewijzen, maar zo'n anderhalf jaar later bleek ik daar toch al een beetje in geslaagd: ik had nog maar net mijn thesisverslag opgepikt of er zat al een mailtje in mijn inbox van ene T.C. met als titel 'proficiat + FWO?' en daarin de uitnodiging om me te begeleiden bij de uitwerking van mijn doctoraatsaanvraag. Met een bereidwillige promotor én een masterdiploma op zak moest ik nu enkel nog het FWO zien te overtuigen van mijn capaciteiten. Een jaar en een motivatiebrief vol "zelfstoef" later was ook dat gelukt: het avontuur kon beginnen. Aan dat doctoraatsavontuur komt met deze thesis een einde. Ik wil van dit dankwoord graag gebruik maken om – met een glimlach op het gezicht – even terug te blikken op de afgelopen vier jaar, die dankzij heel wat mensen zonder twijfel onvergetelijk zijn geworden. Om het met de woorden van Fall Out Boy te zeggen: Thnks fr the mmrs!

Ik begin met het FWO, dat mij in de selectieronde van 2014 heeft uitgekozen als één van de *lucky few*, en me zo de kans – en, misschien belangrijker, de middelen – heeft gegeven om mezelf te ontplooien als taalkundig onderzoeker en dit doctoraat tot een goed einde te brengen.

Natuurlijk heeft ook mijn promotor, Timothy Colleman (ondertussen mag ik gewoon "Timothy" zeggen), daarbij een grote rol gespeeld. Toen ik voor het eerst te horen kreeg dat hij me wou begeleiden bij de projectaanvraag heb ik hem bedankt voor de steun en kansen die hij me bood. Ook vier jaar later wil ik die woorden graag herhalen: bedankt, Timothy, voor het vertrouwen en de waardevolle ondersteuning van de afgelopen jaren. Timothy is, als ik dat mag zeggen, een veeleisende promotor. Hij vindt het bijvoorbeeld

belangrijk dat een onderzoek ook een zekere theoretische relevantie heeft: corpusonderzoek uitvoeren en je resultaten beschrijven, allemaal goed en wel, maar wat leert ons dat nu over het grotere theoretische kader? Als je even vastzit, komt hij ook niet meteen met een kant-en-klare oplossing maar laat hij je er zelf nog even op kauwen. Frustrerend soms, dat wel, maar je leert er ontzettend veel uit. Je moet misschien soms wat tussen de lijntjes leren lezen, maar uit de kleine dingen merk je wel dat Timothy je inzet en kunnen naar waarde schat en het beste in je naar boven probeert te halen. De laatste maanden – wanneer de twijfel al wel eens durfde toe te slaan – waardeerde ik het dan ook dat hij me af en toe verzekerde dat hij er alle vertrouwen in had dat het goed zou komen. En het is ook goed gekomen, hoop ik (beste lezer, oordeel vooral zelf).

Naast Timothy wil ik ook graag alle andere leden en ex-leden van de afdeling Nederlandse Taalkunde bedanken voor de fijne werksfeer, de plezante koffiepauzes en de vele after-work activiteiten. Een aantal collega's wil ik in het bijzonder in de bloemetjes zetten:

... Valerie Bouckaert. Zij bemant de bibliotheek en het secretariaat Nederlandse Taalkunde, maar eigenlijk doet ze zoveel meer dan dat. Zeker in de laatste maanden nam ze me maar al te graag wat werk uit handen door voor mij een boek in te scannen of me een koffie te brengen. Als wederdienst luisterde ik dan ook graag naar de fratsen van haar kat en konijn.

... de *original gang* uit bureau 100.053, Tim en Anne-Sophie. Ik kon bij hen altijd terecht voor een babbel tussendoor of enkele motiverende woorden. Anne-Sophie bedank ik ook graag voor de hulp bij allerlei statistische, technische of lay-outkwesties – iedereen weet dat zij daar een echte crack in is. Tim werd aangesteld als mijn peter en heeft die rol zeker ter harte genomen. Waar de Senseo stond en hoe de kopieermachine werkte, heb ik wel zelf moeten ontdekken (ja, dat zat eigenlijk in jouw takenpakket, Tim), maar verder had ik zeker niet te klagen. Als echte syntax buddies gingen we samen op congres, maar er was ook al eens tijd om een raadseltje op te lossen, een danske te placeren of een Pokémon te vangen ;-).

... de (intussen ex-)collega's uit bureau 100.051, Inge en Steven. Hoewel ik een overtuigd kattenmens ben en Inge een duidelijk hondenmens – iets wat tot verhitte discussies had kunnen leiden – konden we het toch goed met elkaar vinden. Ook met Steven klikte het meteen. We hadden allebei een gezonde dosis zelfrelativeringsvermogen, maar konden ook wel enige appreciatie opbrengen voor de nodige drama en *gossip*.

... de buren bij Taalonthaal. Bedankt voor de ontspannende panfluittonen.

... Lien, de jongste telg bij het Woordenboek van de Vlaamse Dialecten. Net zoals ik een echte foodie (of dat denken we toch), die op de hoogte is van de nieuwste hippe eetplekjes in Gent en omstreken. Geef ons maar een donut van Donuttello of Hoeked, in plaats van Panos. Daarnaast bleken we ook een voorliefde te delen voor tv-series waar we, als we

eerlijk zijn, eigenlijk al te oud voor zijn (denk: doelgroep young adults, of hoe noemen ze dat?).

... de "jonkies" die in oktober 2016 de afdeling Nederlandse Taalkunde kwamen versterken, Amélie en Arne. Aangezien we alle drie Timothy als promotor hadden, doopten we onszelf Team Timothy en maakten we zowel binnen- als buitenland onveilig met onze theoretisch relevante lezingen. Ook op persoonlijk vlak bleken we heel wat gemeenschappelijke interesses te hebben. Bij Arne kon ik altijd gaan aankloppen voor een taalkundig-verantwoorde babbel of een redder in nood, maar ook voor een lekkere kers en een diepzinnig gesprek over alpaca's, katten, Eurosong of bonsaiboompjes. Nadat ik twee jaar van Tim het goede voorbeeld had gekregen, was het nu mijn beurt om mij te bewijzen als meter van Amélie. Ze heeft het soms wat moeilijk met namen (sorry Bert, of was het nu toch Patrick?) en durft 's nachts al eens op zoek te gaan naar *my preciousssss*, maar verder is Amélie een schat van een metekind. We gaven elkaar modeadvies, testten de nieuwe filters op Snapchat uit, gingen levensgevaarlijke tv-opnames bijwonen, zijn ooit eens een Limburger kwijtgeraakt in de gangen van de Blandijn en durfden ook al eens een mondje Duits te praten aan de telefoon. Kortom, we hebben samen heel wat gelachen, meer dan eens tot tranen toe.

Ik wil ook een aantal andere mensen bedanken die op een of andere manier hebben bijgedragen tot mijn onderzoek of mijn doctoraatservaring.

... niet in het minst natuurlijk de leden van mijn doctoraatsbegeleidingscommissie, Jóhanna Barðdal, Muriel Norde en Graeme Trousdale, die van zich bij het begin erg enthousiast toonden over het onderzoek dat ik uitvoerde. Tijdens onze DBC-vergaderingen stonden ze klaar met lovende woorden, nuttige feedback en nieuwe inspirerende ideeën.

... Steven Claeyssens van de KB en Guy De Pauw van Textgain, zonder wie het niet mogelijk was geweest om gebruik te maken van dat immense Delphercorpus.

... alle mededoctoraatsstudenten in binnen- en buitenland, dankzij wie de vele congressen en workshops niet *all work,* maar ook een beetje *play* waren.

... Gitte Callaert, voor haar oog voor detail en lay-outtips tijdens de allerlaatste eindsprint.

Het mag duidelijk zijn dat ik mijn doctoraat vooral heb "beleefd" (of "overleefd") op de universiteit. Uiteraard wil ik ook nog een aantal andere mensen bedanken die me van buitenaf hebben gesteund en die het vooral mogelijk maakten om dat doctoraat ook af en toe eens aan de kant te schuiven.

... Lise, Lesley, Niek, Steffie, de ex-medestudenten van MTB, en anderen die ik hier vergeet. Bedankt voor de vele leuke Messengergesprekken, Gent- of Antwerpendates, Bruglunches, feestjes, terrasjes, gezellige treinritten, enz. Ook dank aan de werkman-met-blauwe-trui, die met zijn vrolijke dansjes en enthousiast gezwaai het uitzicht op de

bouwwerf van de Boekentoren en de geluidsoverlast die daarbij kwam kijken toch wat draaglijker maakte.

... alle medecursisten van Japans (het Perspectief): ありがとうございます! Een keer per week was het eens volledig *back to basics*. Gaandeweg ontdekken hoe een taal als het Japans volledig anders in elkaar zit dan het Nederlands was voor een taalnerd als ik natuurlijk pure ontspanning.

... mijn schoonfamilie. Hoewel ze vermoedelijk maar een vaag idee hadden van waar ik nu eigenlijk een hele dag mee bezig was, bleven ze interesse tonen in mijn doctoraat en vroegen ze af en toe of alles naar wens verliep. Dat apprecieer ik! Het was ook altijd tof om bij hen op bezoek te gaan en mijn zinnen eens te verzetten tijdens een van de vele cinema-uitjes, barbecues, koffievisites of andere eet- en drinkpartijtjes.

Eindigen doe ik met de belangrijkste mensen in mijn leven.

... mijn ouders, die me van meet af aan hebben gesteund in alles wat ik tot nu toe al heb ondernomen in mijn leven. Ik heb altijd het gevoel gehad dat ze trots op me waren, en dat doet deugd! Dankzij hen heb ik geleerd om voor mezelf de lat hoog te leggen en altijd naar de beste versie van mezelf te streven, en dat heeft me zeker geholpen om dit doctoraat tot een goed einde te brengen.

... mijn vriend, Bono, zou ik voor véél dingen kunnen bedanken, maar dan vooral voor zijn geduld. Ja, het lijkt wat cliché. Ik denk dat zowat iedereen hun partner bedankt voor het geduld dat ze hebben moeten opbrengen tijdens die laatste maanden van het doctoraat. Maar geloof me vrij dat dat in zijn geval zeker geen eenvoudige opgave kan zijn geweest. Als ik eens een mindere dag had gehad op de Blandijn, durfde ik die frustraties wel eens op hem uitwerken of was het de hele avond “beeld zonder klank”. Ik kan niet beloven dat al mijn kuren zullen verdwijnen nu mijn doctoraat is ingediend, maar als je een Doctor als vriendin wilt, moet je dat er maar bijnemen, hé (*kidding, kind of*).

Zo, voor mij zit het er bijna op, maar voor u, beste lezer, begint het natuurlijk nog maar pas. Ik zou u dan ook graag veel leesplezier wensen, al ben ik me ervan bewust dat velen niet veel verder dan dit dankwoord zullen geraken. Toch hoop ik dat er voldoende mensen zijn die, net zoals mij, geïntrigeerd zijn door het feit dat we ons blijkbaar wel een hoedje of een aap kunnen schrikken, maar zelden of nooit een mutsje of een giraf. Ook wie zich al nachtenlang heeft suf gepiekerd over productiviteit en constructionele netwerken, leest maar beter verder. Ik beloof dat u er geen spijt van zult krijgen.

Emmeline Gyselinck
Gent, september 2018

List of Translations

This is an alphabetical list of all Dutch intensifiers and verbs that are part of the synchronic and diachronic data sets, with their near-literal translations in English. Words marked with an asterisk are fictitious diseases, most of which are in a way derived from an actual disease; words marked with ? are nonsense words that are hard to translate.

Intensifiers

[n°] slagen in de rondte	[n°] punches around
adellijk blauw	royal blue
apelazerus	monkey-lazarus* (~leprous)
azuur-blauw	azure-blue
beroerd	miserable
beten en scheten	bites and farts
bewusteloos	unconscious
bicblauw	pen-blue
blaren	blisters
blaren op de tong	blisters on the tongue
blauw	blue
blauw en groen	blue and green
blauw en paars	blue and purple
bleek	pale
blind	blind
bloot	naked
bont en blauw	black and blue
de adem uit de longen	the breath out of the lungs
de balg uit het lijf	the gut out of the body
de benen PREP het gat	the legs out of the butt
de benen PREP het lijf	the legs out of the body
de benen uit de naad	the legs out of the seam
de benen uit het lid	the legs out of the joint
de blaren	the blisters
de blaren in de handen	the blisters in the hands
de blaren op de hakken	the blisters on the heels
de blaren op de tond [sic]	the blisters on the ?
de blaren op de tong	the blisters on the tongue
de blaren PREP de voeten	the blisters on the feet
de blaren op de zolen	the blisters on the soles
de blaren op het verhemelte	the blisters on the palate
de blubber	the blubber
de bril van het hoofd	the glasses off the head
de buik rond	the belly round
de griebels	the shivers
de hakken scheef	the heels crooked
de handen blauw	the hands blue
de handen kapot	the hands broken
de handen stuk	the hands broken
de hersenen suf	the brains drowsy
de heupen stuk	the hips broken
de hik	the hiccups
de klere	the cholera

de kleren van het lijf	the clothes off the body
de kolere	the cholera
de krampen	the cramps
de ledematen uit de gewrichten	the limbs out of the joints
de longen leeg	the lungs empty
de longen stuk	the lungs broken
de longen te barsten	the lungs to bursts
de longen uit de balg	the lungs out of the gut
de longen uit het lijf	the lungs out of the body
de mazelen	the measles
de naad uit de broek	the seam out of the pants
de naad uit het lijf	the seam out of the body
de nagels blauw	the nails blue
de nieren los	the kidneys loose
de ogen blind	the eyes blind
de ogen uit de kassen	the eyes out of the sockets
de ogen uit het hoofd	the eyes out of the head
de ogen zat	the eyes drunk
de oren rood	the ears red
de oren van het hoofd	the ears off the head
de pest	the plague
de pestpokken	the plague-smallpox
de pleuris	the pleurisy
de pleuris uit het lijf	the pleurisy out of the body
de pokken	the smallpox
de poten kapot	the legs broken
de poten PREP het lijf	the legs out of the body
de poten van onder de keukenstoel	the legs from under the kitchen table
de rambam	?
de schoenen vanonder hun voeten	the shoes from under the feet
de stuipen	the fits
de stuipjes	the little fits
de takken	the piles (colloquial term)
de tering	the consumption
de tering-takke	the consumption-piles
de tranen	the tears
de typhus	the typhoid
de vingers beurs	the fingers mushy
de vingers blauw	the fingers blue
de vingers groezelig	the fingers grubby
de vingers krom	the fingers bent
de vingers moe	the fingers tired
de vingers wond en rond	the fingers sore and round
de vinketering	the finch-consumption
de voeten PREP het lijf	the feet out of the body
de zenuwen	the nerves
de ziel dood	the soul dead
de ziel uit de naad	the soul out of the seam
de ziel uit de raap	the soul out of the head
de ziel uit het lijf	the soul out of the body
de zolen PREP de voeten	the soles off the feet
de zolen uit de sloffen	the soles off the slippers
de zolen van de schoenen	the soles off the shoes
donkerblauw	dark blue
dood	dead
een aanp [sic]	?
een aap	a monkey
een barst	a crack
een beroerte	a stroke
een blauw hart	a blue heart
een breuk	a fracture
een bult	a hump
een delirium	a delirium
een delirium tremens	a delirium tremens
een deuk	a fit
een eind in de rondte	a distance around
een halve beroerte	a half stroke
een hart in het lijf	a heart in the body
een hartaanval	a heart attack
een hartverlamming	a heart paralysis
een hoed	a hat
een hoedje	a little hat
een houten hart	a wooden heart
een kokosnoot	a coconut
een koliek	a colic
een kontzweer	(a) haemorrhoid(s)
een kriek	a hump
een liesbreuk	a groin hernia
een loei	a whopper
een mik	a belly

een ongeluk	an accident
een ootje	?
een pissebed	an isopod/sow bug
een puist	a pimple
een pukkel	a pimple
een punthoofd	a pointy head
een rolberoerte	a fit
een rotje	a cracker
een slag in de rondte	a punch around
een slaghoedje	a percussion cap
een stuip	a fit/spasm
een stuk in de gilet	a piece in the cardigan
een stuk in de hakken	a piece in the heels
een stuk in de kont	a piece in the butt
een stuk in de kraag	a piece in the collar
een stuk in de voeten	a piece in the feet
een zoeavenmuts	a zouave-bonnet
een zuurstok	a stick of rock
flauw	faint
gaar	cooked
geel	yellow
geel en groen	yellow and green
gek	crazy
grasgroen	grass green
grijs	grey
grijs en groen	grey and green
groen	green
groen en blauw	green and blue
groen en geel	green and yellow
halfdood	half dead
halfgek	half crazy
halfkapot	half broken
halflam	half lame
halfslap	half weak
halfsuf	half drowsy
halfziek	half sick
het apelazerus	the monkey-leprosy*
het apenzweet	the monkey-sweat*
het apezuur	the monkey-heartburn*
het geel	the yellow
het hart uit het lijf	the heart out of the body
het hoedje	the little hat
het hoofd gek	the head crazy
het hoofd suf	the head drowsy
het laplazerus	the lap-leprosy*
het lazerus	the leprosy
het leplazerus	the lep-leprosy*
het licht uit	the light out
het licht uit de ogen	the light out of the eyes
het ongans	the biliousness
het pleuris	the pleurisy
het rambam	?
het schompes	the schompes*
het schuim op de hiel	the foam on the heel
het schuim op de mond	the foam on the mouth
het schuim op de ziel	the foam on the soul
het snot voor de ogen	the snot before the eyes
het vel van de botten	the skin off the bones
het vuur uit de molières	the fire out of the lace-ups
het vuur uit de pen	the fire out of the pen
het vuur uit de rennerssloffen	the fire out of the cycling slippers
het vuur uit de schaatsen	the fire out of the skates
het vuur uit de schenen	the fire out of the shins
het vuur uit de schoenen	the fire out of the shoes
het vuur uit de sloefen	the fire out of the slippers
het vuur uit de sloffen	the fire out of the slippers
het vuur uit de slofjes	the fire out of the little slippers
het vuur uit de slofkens	the fire out of the little slippers
het vuur uit de sokken	the fire out of the socks
het vuur uit de spaken	the fire out of the spokes
het vuur uit de spikes	the fire out of the spikes
het vuur uit de sportschoenen	the fire out of the trainers
het vuur uit de sportsloffen	the fire out of the sports slippers
het vuur uit de vingers	the fire out of the fingers
het zuur	the heartburn
in de poeier	in the powder (~ shattered)
in het zweet	in the sweat
in pust	in pimple
kapot	broken

klem	stuck/drunk/stuffed
kleurenblind	colour blind
krampen	cramps
krankjorum	bonkers
kreupel	crippled
krom	bent
lam	lame
laveloos	sloshed
lazerus	leprous/sloshed
ledematen blauw	limbs blue
leeg	empty
lens	weak
murw	mellow
ongans	unwell
ongelukkig	unhappy
onnozel	silly
over de kop	over the head
paars	purple
paars en groen	purple and green
plat	flat
pleuris	pleurisy
rond	round
rood en groen	red and green
rot	rotten
scheef	crooked
scheel	cross-eyed
schor	hoarse
slap	weak
spinaziegroen	spinach green
stom	stupid
stuipen	fits
stuk	broken
suf	drowsy
te barsten	to bursts
te blubber	to blubber
te pletter	to smithereens
te pleuris	to pleurisy
te sappel	to worries (sappelen = to be worried)
ten doode	to death
tranen	tears
tureluurs	crazy
uit de naad	out of the seam
uit de naden	out of the seams
uit het lid	out of the joint
uit het lood	out of the lead (~off one's balance)
verloren	lost
verrot	rotten
VLD-blauw	VLD-blue
wezenloos	vacant/blank
wild	wild
witjes	white-ish
zenuwziek	neurotic
ziek (en weer gezond)	sick (and healthy again)
zot	crazy
zwart	black

Verbs

aaien	to stroke
acteren	to act
adverteren	to advertise
analyseren	to analyse
annonceren	to announce
applaudisseren	to applaud
argumenteren	to argue
associëren	to associate
babbelen	to chatter
baden	to bathe
baggeren	to dredge
bakken	to bake
balanceren	to balance
balen	to be fed up with
bellen	to call
beminnen	to love
besparen	to economise
betalen	to pay
betogen	to demonstrate
beuken	to batter
bewapenen	to arm
bezetten	to occupy
bezuinigen	to economise
bibberen	to shiver
bidden	to pray
bladeren	to leaf through
blaffen	to bark
blazen	to blow

bloeien	to blossom
blokken	to study/cram
blowen	to smoke weed
boenen	to polish
boetseren	to mould
borduren	to embroider
boren	to drill
borstelen	to brush
bouwen	to build
breien	to knit
broeden	to brood
brullen	to roar
buigen	to bow
bula'en [sic]	to live
chatten	to chat
cijferen	to make calculations
citeren	to cite
combineren	to combine
communiceren	to communicate
compromitteren	to compromise
concurreren	to compete
confereren	to confer
congresseren	to hold a conference
consumeren	to consumate
controleren	to control
creëren	to create
dansen	to dance
debatteren	to debate
demarreren	to break away
demonstreren	to demonstrate
denken	to think
dichten	to write poetry
dirigeren	to conduct
discussiëren	to discuss
dobbelen	to dice
dobberen	to float
dokteren	to be in practise
doperen	to take dope
downloaden	to download
draaien	to turn
dragen	to carry
draven	to trot
dreigen	to threaten
dresseren	to train
drinken	to drink
drogeren	to take drugs
dromen	to dream
drukken	to press
dubben	to dub
duiken	to dive
duwen	to push
e-mailen	to email
eten	to eat
experimenteren	to experiment
feesten	to party
fietsen	to cycle
filmen	to film
filosoferen	to philosophise
fingeren	to feign
fladderen	to flutter
fluiten	to whistle
foeteren	to grumble
forceren	to force
fotograferen	to take photographs
fuiven	to party
gamen	to game
gapen	to gape
gebruiken	to use
geeuwen	to yawn
genieten	to enjoy
giechelen	to giggle
gieren	to shriek
gillen	to screech
gluren	to peek
gniffelen	to snigger
gokken	to gamble
golfen	to (play) golf
googlen	to google
gooien	to throw
grabbelen	to scramble
grappen	to joke
graven	to dig
grijnzen	to grin
groeien	to grow
handelen	to trade
hijgen	to pant
hinniken	to neigh
hoesten	to cough
hollen	to run
hongeren	to hunger
huilen	to cry
huren	to hire

ideologiseren	to ideologise
internetten	to surf the Net
investeren	to invest
isoleren	to isolate
jagen	to hunt
jakkeren	to slave away
janken	to howl
jazzen	to play jazz
joggen	to jog
juichen	to cheer
kaatsen	to bounce
kakelen	to cackle
kakken	to poop
kandideren	to put oneself up for
kappen	to chop
kiezen	to choose
kijken	to look
klagen	to complain
klappen	to clap
kletsen	to chatter
klikken	to click
kloppen	to knock
knagen	to gnaw
kniezen	to mope
knijpen	to squeeze
knippen	to cut
knipperen	to blink
knokken	to fight
knuffelen	to cuddle
koersen	to race
koken	to cook
kopen	to buy
kotsen	to puke
krabben	to scratch
kreunen	to moan
kwellen	to torment
lachen	to laugh
leggen	to lay
lenen	to borrow
leren	to learn
leuteren	to drivel
leven	to live
lezen	to read
liegen	to lie
liften	to hitchhike
lijnen	to diet
lobbyen	to lobby
lonken	to ogle
lopen	to run
lullen	to (talk) bullshit
mailen	to mail
manipuleren	to manipulate
manoeuvreren	to manoeuvre
mediteren	to meditate
mekkeren	to bleat
meppen	to smack
mijmeren	to muse
molenwieken	to wing
musterberen [sic]	to ?
naaien	to sew
neuken	to fuck
niezen	to sneeze
nuanceren	to nuance
oefenen	to practise
onderhandelen	to negotiate
organiseren	to organise
orkestreren	to orchestrate
overleggen	to confer
paaien	to mate
pachten	to lease
paffen	to puff
pakken	to take
patrouilleren	to patrol
peilen	to gauge
peinzen	to ponder
pendelen	to commute
pennen	to scribble
persen	to strain
pezen	to slave away
piekeren	to worry
pijnigen	to torture
pingelen	to haggle
plannen	to plan
playbacken	to lip-sync
pleiten	to plead
ploeteren	to plod
poetsen	to clean
pompen	to pump
prakkiseren	to brood
praten	to talk
preken	to preach
presenteren	to present

prijzen	to prize
prikken	to inject
printen	to print
procederen	to litigate
produceren	to produce
programmeren	to program
protesteren	to protest
prutsen	to fiddle
puzzelen	to puzzle
racen	to race
raden	to guess
rappen	to rap
ravotten	to romp
recenseren	to review
recyclen	to recycle
redeneren	to reason
regelen	to regulate
regeren	to rule
registreren	to register
reizen	to travel
rekenen	to calculate
rekken	to stretch
relativeren	to put into perspective
rennen	to run
reorganiseren	to reorganise
repeteren	to rehearse
reserveren	to reserve
rijden	to drive/ride
roddelen	to gossip
roeien	to row
roepen	to yell
roeren	to stir
roffelen	to ruffle
roken	to smoke
rollen	to roll
sakkeren	to grumble
samplen	to sample
sappelen	to worry
scanderen	to chant
schaatsen	to skate
schakelen	to change gear
scheiden	to separate
scheppen	to shovel
scheuren	to tear
schieten	to shoot
schijnen	to shine
schilderen	to paint
schitteren	to shine
schminken	to make up
schnabbelen	to have a job on the side
schrappen	to work
schreeuwen	to scream
schreien	to weep
schreppen [sic]	to ?
schrijven	to write
schrikken	to be startled
schuimen	to foam
schuiven	to slide
schuren	to chafe
scrabbelen	to play scrabble
selecteren	to select
serveren	to serve
shoppen	to shop
signalen [sic]	to signal
sikkeneuren	to nag
sjouwen	to drag
slaan	to hit
slapen	to sleep
slempen	to feast
slepen	to lug
sleuren	to haul
sleutelen	to tinker with
slikken	to swallow (drugs)
sloffen	to shuffle
sloven	to drudge
smeren	to smear
smokkelen	to smuggle
sms'en	to text
snikken	to sob
snoeien	to trim
snoepen	to eat sweets
snoeven	to swagger
snuiten	to blow one's nose
snuiven	to inhale (drugs)
solliciteren	to apply for a job
sparen	to save
spartelen	to flounder
spelen	to play
speuren	to investigate
sponsoren	to sponsor

sporten	to sport
spreken	to speak
springen	to jump
sprinten	to sprint
spuiten	to inject (drugs)
spurten	to sprint
staken	to strike
stampen	to stomp
staren	to stare
steken	to stab
stelen	to steal
stoken	to heat
stomen	to steam
storen	to disturb
storten	to crash
stoten	to bump
strelen	to caress
strijken	to iron
studeren	to study
sturen	to send
suffen	to drowse
supporteren	to cheer for
surfen	to surf
swingen	to swing
tappen	to tap
tekenen	to draw
telefoneren	to call
telegraferen	to telegraph
tellen	to count
tennissen	to play tennis
tikken	to type
tillen	to lift
time-managen	to manage time
tobben	to fret
toeren	to tour
toeteren	to honk
tollen	to spin
tongzoenen	to french kiss
trainen	to train
transformeren	to transform
transpireren	to perspire
trappen	to pedal
trekken	to pull
treuren	to sorrow
trommelen	to drum
trompetteren	to trumpet
turen	to peer
turnen	to practise gymnastics
turven	to tally
uitleggen	to explain
vallen	to fall
varen	to sail
vasten	to fast
vechten	to fight
vegen	to sweep
verdienen	to earn
vergaderen	to meet
vergelijken	to compare
verkopen	to sell
vernieuwen	to renew
verschieten	to be startled
verschrikken	to be frightened
verzamelen	to collect
verzinnen	to invent
vliegen	to fly
vloeken	to curse
voetballen	to play soccer
vragen	to ask
vreten	to gorge oneself
vrijen	to make love
waarschuwen	to warn
wachten	to wait
wankelen	to stagger
wassen	to wash
wegen	to weigh
wenen	to cry
werken	to work
werpen	to throw
werven	to hire
wiebelen	to wiggle
winkelen	to shop
wisselen	to change
worstelen	to wrestle
wrijven	to rub
wringen	to wring
wroeten	to root
zagen	to nag
zappen	to zap
zeilen	to sail
zeulen	to lug
zeuren	to nag

zich amuseren	to enjoy oneself
zich enerveren	to be agitated
zich ergeren	to be annoyed
zich generen	to be embarrassed
zich integreren	to integrate oneself
zich isoleren	to isolate oneself
zich schamen	to be embarrassed
zich verbazen	to be amazed
zich verheugen	to rejoice
zich vermaken	to have fun
zich vervelen	to be bored
zien	to see
zingen	to sing
zitten	to sit
zoeken	to search
zuigen	to suck
zuipen	to booze
zwaaien	to wave
zwemmen	to swim
zweten	to sweat
zwijgen	to shut up
zwoegen	to labour

List of Tables

List of Figures

Table of Contents

Chapter 1
Introduction

This doctoral dissertation is aimed at elucidating the factors and mechanisms that underlie shifts in the internal organisation of the constructional network. Building on the theory of Diachronic Construction Grammar, we set out to demonstrate how quantitative and qualitative changes in the use of a construction relate to changes in productivity and schematicity at different levels of abstraction in the constructional network hierarchy. Concretely, we will investigate the synchronic and diachronic variation attested in one specific construction in Dutch, focusing on national variation between Belgian and Netherlandic Dutch in present day and the recent diachronic developments (19th-21st Century) of the construction in Netherlandic Dutch. The intensifying fake reflexive resultative construction is illustrated in the following examples (see also §1.1).

(1) Overtreders, mensen die beboetbaar bezig zijn geweest, **schrikken zich** vaak **een hoedje** over de hoogte van de boete. (SoNaR)
[...] startle themselves often a little hat [...]
'Offenders, people who have committed a punishable offence, are often highly startled by the amount of the fine.'

(2) Ik zapte er nu langs en **schrok me de blaren van het geplamuurde gezicht** met die rode lippen. (Twitter, 23/10/2016)
[...] startled myself the blisters off the plastered face [...]
'I swooped by and was highly startled by those red lips.'

The intensifying fake reflexive resultative construction presents a number of features that make it a suitable subject for a corpus-based study into (recent) shifts in productivity and the reorganisation of the constructional network. First of all, the intensifying fake reflexive resultative construction appears to be highly productive in present-day Dutch: the sets of both verbs and intensifiers that may appear in this construction display a lot of syntactic and semantic variability (cf. §1.1 infra). However, while the degree of variability observed might at first glance create the impression that "anything goes", closer inspection reveals that there is also a considerable degree of conventionality

involved in the use of this construction. With regard to the above examples, native speakers of Dutch will in all likelihood concur that (1) is a much more conventional way of saying that one is highly startled than (2). This raises the question why some lexical items are frequently recruited by the construction as (conventional) intensifiers, whereas other potential candidates are barely picked up at all. This interaction between creative innovation and conventional conservatism does not only apply to the individual verb and intensifier slots but also to their mutual combinatorics. While some verbs and intensifiers display considerable flexibility in their combinatorial behaviour, others are confined to a much more limited set of collocates. This suggests that there are certain conventions with respect to the verb-intensifier combinations that speakers of Dutch will be inclined (or disinclined) to use.

Second, the present-day productivity and flexibility of the construction appears to be a rather recent phenomenon, which makes it possible to trace this development in digitally available corpora. Judging by the citations in the *Woordenboek der Nederlandsche Taal* ['Dictionary of the Dutch Language', in what follows: WNT] and some examples in the Corpus Literair Nieuwnederlands ('Corpus of literary Modern Dutch', Geleyn 2016), the earliest attestations of the intensifying fake reflexive resultative construction mainly featured the intensifier *dood* 'dead' (see §1.2). There are indications that the construction has undergone a massive expansion since the early 19th Century, both in terms of its frequency of use and with respect to the variety of verbs and intensifiers it hosts. This recent innovation is likely related to the meaning component of the construction. It has been argued that the linguistic domain of intensification is characterised by a need for expressivity, and that the linguistic means used to express this intensification are subject to rapid change, innovation and renewal. While a detailed description of the construction will be provided in Chapter 2, we will already briefly introduce the construction and the variation that is attested in present-day Dutch in the next section (§1.1). The aims of this thesis and the concrete research questions will be discussed in more detail in the second section (§1.2). The final section of this introductory chapter presents an outline of the thesis (§1.3).

1.1 A first introduction to the Dutch intensifying fake reflexive resultative construction

While the intensifying fake reflexive resultative construction has the same syntactic structure as the regular (i.e. literal) fake reflexive resultative construction, viz. [SUBJ V REFL XP], it does not carry the same resultative semantics. Thus, there is no obvious

structural difference between examples (3) and (4) below, but there is a clear difference in the semantic contribution of the element *dood* 'dead' in the two examples and, accordingly, in the respective meanings of the entire clauses.

(3) De man dronk zich dood op vroege leeftijd.
the man drunk himself dead [...]
'The man had drunk himself to death at an early age.'

(4) De man schrok zich dood toen hij de muis zag.
the man startled himself dead [...]
'The man was highly startled when he saw the mouse.'

In the former example, the adjective *dood* 'dead' is a true resultative phrase that denotes an actual result of the verbal activity, whereas in the latter it functions as an intensifier, indicating that the verbal activity is boosted or performed with a heightened intensity. As we will show in §2.2.2, the use of some kind of formally resultative pattern in order to convey an intensifying meaning is not unique to Dutch, but Dutch sets itself apart from other languages by the range of verbs and, especially, the variety of different intensifiers that can be used in the construction. The examples below illustrate a mere fraction of the numerous possibilities in present-day Dutch (taken from the SoNaR corpus).

(5) Met het CDA ergeren ik en velen met mij zich **groen en geel** aan de graffititerreur van de jeugd die in veel gevallen niet eens foutloos kan spellen.
[...] annoy I and many with me themselves green and yellow [...]
'Like the CDA, I and many others are very annoyed by the graffiti terror of the adolescents, whose spelling is often far from flawless.'

(6) Ze liegen zich **te barsten**, net als de doorsneebezoeker aan een goktent in Las Vegas.
[...] they lie themselves to bursts [...]
'They are lying their butts off, just like the average visitor in a gambling den in Las Vegas.'

(7) De vrouw schrok zich **een aap** toen na de relatie bleek dat ze een lening moest betalen die C.C. met haar vervalste handtekening had afgesloten.
the woman startled herself a monkey [...]
'The woman was very startled when it turned out that she had to pay off a loan that C. C. had contracted by falsifying her signature.'

(8) Een vijftiental beloften loopt zich **de pleuris uit het lijf**; Gert Verheyen (36) kijkt goedkeurend toe.
some fifteen reserves run themselves the pleurisy out of the body [...]
'Some fifteen reserves are running themselves to pieces; Gert Verheyen (36) is watching in approval.'

Not only can the intensifiers take different syntactic forms (e.g. AP, PP, NP or NP+PP), they can also be recruited from multiple semantic domains. It will be shown in §2.3 that intensifiers often have negatively connoted original semantics– which is indeed to a greater or lesser extent the case for several of the intensifiers in the above examples, e.g.

dood 'dead', *de pleuris uit het lijf* 'the pleurisy out of the body' and *te barsten* 'to bursts' –, but Dutch has also developed peculiar intensifiers, the origins of which are much less obvious (e.g. *een hoedje* 'a little hat'). It was already mentioned above that this construction appears to be quite productive in present-day Dutch. As is illustrated by (2) and other examples from Twitter below, the intensifying fake reflexive resultative construction allows for some witty or creative choices in its intensifier slot.

(9) Paps bed aan het maken, net doorheen gezakt god schrok me **de tieten van me lijf af**. (25/02/2013)
[...] startled myself the tits off my body off
'Dad was fixing my bed and fell through. God, it startled the hell out of me.'

(10) Ik verveel me **de neten** en kan wel gaan leren maar daar heb ik helemaal geen zin in. (19/06/2017)
[...] I bore myself the nits [...]
'I am so bored... I could go and study but I don't feel like it at all.'

(11) Ik was denk ik vergeten dat Florian er stond.... ik schrok me **de knetters** van een karton. (15/08/2017)
[...] startled myself the sparks [...]
'I must have forgotten Florian was there... I was so startled by the piece of cart board.'

At the same time, there are collocational preferences and conventional combinations that keep the linguistic creativity within bounds. The intensifier *een hoedje* 'a little hat', for instance, is almost exclusively used with the verb *schrikken* 'to be startled' and *groen en geel* 'green and yellow' enters into a fixed collocation with *zich ergeren* 'to be annoyed'. As we will show, the differences in collocational behaviour translate to varying degrees of productivity at different levels in the hierarchy of the constructional network. From a diachronic perspective, changes in the collocational preferences and productivity of these patterns trigger certain shifts and internal reorganisations in the network, which are at the centre of this investigation. The next section sets out the precise research aims and formulates some concrete questions that we aim to provide an answer to throughout the thesis.

1.2 Aims of this thesis

The research presented in this thesis focuses on synchronic variation and diachronic shifts in the intensifying fake reflexive resultative construction since the early 19th Century. The present-day situation that was briefly illustrated in the previous section appears to be the result of a rather recent expansion. Although some examples were already attested before 1800, there are clear indications that the construction had not

really taken off yet: most pre-19th Century examples that we could find in the WNT feature the adjective *dood* 'dead', suggesting that the intensifier slot was still highly constrained at the time (see §2.2.2.3 for more details). Based on these observations, we expect to find a series of changes in the internal structure of the constructional network of the intensifying fake reflexive resultative construction over the past 200 years or so. The aim of this thesis is not to provide an account of the entire history of this construction, but to investigate the changes it has undergone in its recent history, viz. since the beginning of the 19th Century. In that regard, this work aligns with a current research line in Diachronic Construction Grammar that is not primarily concerned with the emergence of new constructions (i.e. constructionalisation, in terms of Traugott & Trousdale 2013), but with the way in which constructions continue to undergo changes after they have become established. In order to provide an account of the expansion and constructional changes that the intensifying fake reflexive resultative construction seems to have undergone, we zoom in on both large and small-scale shifts within the constructional network. We will investigate whether the corpus data allow us to construct a possible representation of the intricate structure of the constructional network, and elucidate how this structure might have changed over the past two centuries.

The investigation starts out from the use of the intensifying fake reflexive resultative construction in present-day Dutch. A number of research questions can be formulated.

I. General use: What is the overall frequency of this construction in our corpus? What kind of variation do we find in the verb and intensifier slots of the construction? Are some verbs and intensifiers better represented than others?

II. Specific verb-intensifiers combinations: Do some intensifiers show clear preferences with respect to the verbs they occur with, or, vice versa, do certain verbs exhibit important selectivity in the intensifiers that may be used to boost the verbal activity? How can we account for such collocational preferences or dispreferences?

III. Productivity: How productive are the verb and intensifier slots of the construction overall and at lower levels in the network? What kind of factors play a role in determining the productivity of (sub)schemas at different levels of abstraction?

IV. Constructional networks: What does it mean to build a hierarchically organised constructional network and what is the cognitive reality of such a taxonomy? Which subschemas and micro-constructions need to be represented and at which level are they situated? How are the different nodes in the network interrelated and motivated?

V. Geographical variation: Are there any differences with respect to overall frequency and productivity of the construction between Belgian and Netherlandic Dutch? Do speakers of Belgian and Netherlandic Dutch show varying preferences with respect to the use of certain intensifiers and intensifier-verb combinations? Do we need different network representations to account for such national variation?

In the second part of the study, we will investigate the recent history of this construction in order to identify the changes it has undergone. As was mentioned earlier, the intensifying fake reflexive resultative construction is a promising object of study for such a relatively short-term diachronic investigation because of its expressive meaning: the lifecycle of intensifiers and other expressive forms is typically found to be rather short (Stoffel 1901, Bolinger 1972, Partington 1993, Lorenz 2002, De Clerck & Colleman 2013, inter alia). The following questions will be discussed.

I. General development: How can we characterise the frequency development that the construction has undergone over the past two centuries? Has the construction widened its semantic scope over time? At what time were the verbs and intensifiers that are used in present-day Dutch introduced in the construction? What is the role of expressivity in the creation of new intensifiers? Do individual verbs and intensifiers display different pathways of change or do they develop in largely parallel fashions?

II. Specific verb-intensifier combinations: Are there striking diachronic shifts in the combinatorial preferences of verbs and intensifiers? Do we find evidence of emerging or obsolescing conventional combinations or fixed collocations?

III. Productivity: Has the construction overall become more productive over time? What kind of productivity shifts can be observed at lower levels of abstraction?

IV. Constructional networks: What did the organisation of the constructional network look like in earlier stages? Which reconfigurations have taken place how can these internal shifts be interpreted in terms of changes in productivity and schematicity?

The research questions will be addressed on the basis of large data sets constructed on the basis of two journalistic corpora. By combining present-day newspaper data from the STEVIN Dutch Reference corpus (SoNaR) and historical newspaper data from the Delpher database, we are able to span the entire period from 1830 until 2011. The present-day data contain both Belgian and Netherlandic Dutch, the diachronic corpus is limited to Netherlandic Dutch (see Chapter 3 for more detailed information on this corpus). A

number of quantitative techniques, which will be presented in Chapters 4 and 5, will be applied to the data. Each of these techniques is geared to one particular aspect of the analysis; together, they provide us with the necessary tools to construct a comprehensive and detailed picture of the recent history of the intensifying fake reflexive resultative construction.

We believe that the results of this longitudinal investigation will contribute to the theoretical framework of Diachronic Construction Grammar and, in particular, to our understanding of the hierarchy of constructional networks and the underlying mechanisms that shape and reshape the internal organisation of those networks. In addition, the role of expressivity as a driving force in language change, which is well-documented in the development of individual intensifiers, will be tested against a type of expressive construction that has not received much attention yet in the existing literature.

1.3 Outline of the thesis

In Chapter 2, we will lay out the theoretical groundwork for the investigation. After a brief introduction to Construction Grammar and the recently developed interest in variation and change in constructions, special attention is paid to how the theoretical concepts of productivity and constructional networks are to be interpreted within (Diachronic) Construction Grammar. In the second section, we give a more detailed description of the intensifying fake reflexive resultative construction that was briefly introduced in §1.1. The Dutch construction is situated within the wider context of resultative constructions and a brief survey of the existing literature on this construction is provided. The third section discusses the role of expressivity and intensification in language change. Chapter 3 deals with the corpus and methodology used in this investigation. It motivates why we have opted for journalistic data and introduces the existing journalistic corpora that are at the basis of our larger, continuous corpus. In addition, we give a stepwise, detailed explanation of the compilation and annotation of both the synchronic and diachronic data sets. The analysis of the synchronic data is presented in Chapter 4. Following the research aims in §1.2, the chapter is divided into four main sections, each building on the results of the previous section in order to dig deeper into the use of the construction: (i) general use and frequency, (ii) collocational patterns, (iii) productivity and (iv) constructional networks. Each section is further subdivided in two subsections: the first subsection presents the synchronic use of the intensifying fake reflexive resultative construction in Netherlandic Dutch, the second subsection looks into synchronic variation by comparing the Netherlandic Dutch data to

data from Belgian Dutch. Chapter 5 tracks the diachronic development of the intensifying fake reflexive resultative construction in Netherlandic Dutch. In parallel to Chapter 4, the chapter is divided in four sections focusing on (i) general development, (ii) collocational expansion and conventionalisation, (iii) shifts in productivity and (iv) reorganisation of the constructional network. The most important findings of this investigation are summarised in the first section of Chapter 6. In the second section of Chapter 6, we discuss how our observations tie in with existing research and contribute new insights to the theoretical framework of Diachronic Construction Grammar. Finally, we provide some directions for further research.

Chapter 2
Theoretical preliminaries

This chapter sets the stage for the analyses in the following chapters by introducing a number of topics and concepts which are of particular interest for our research aims and which will continue to be relevant throughout the entire study. First, the theoretical framework will be outlined and it will be discussed how the current investigation aims to contribute to the further development of concepts and ideas established within the domain of (Diachronic) Construction Grammar. After that, we will elaborate on the intensifying fake reflexive resultative construction that was briefly introduced in the previous chapter, specifically focusing on the aspects of the construction which have (or have not) been addressed in the existing literature. It will be illustrated how this construction relates to the more general resultative construction, the use of which is well-documented in several linguistic traditions, and why this construction in particular is such a suitable candidate for the aims of this investigation. The unique character of the intensifying fake reflexive resultative construction will be further motivated in the third section, which focuses on the notions of linguistic expressivity and intensification and their roles in language change.

2.1 A framework for tracking diachronic constructional changes

The study of morphosyntactic change has been dominated by the grammaticalisation framework for several decades (for some fundamental publications on the principles of grammaticalisation, see Meillet 1912, Kuryłowicz 1975, Hopper & Traugott 2003). Grammaticalisation research has yielded a number of detailed empirical case studies of historical changes in many languages and has certainly deepened our understanding of

grammatical change (see Narrog & Heine 2011, inter alia). Although some of the basic principles of grammaticalisation have recently been called into question (see, e.g., Campbell 2001, Janda 2001, Joseph 2001, Norde 2009 for some key arguments in this debate), it is still a well-established framework in historical linguistics and many current studies at least partly rely on existing grammaticalisation models. While the earliest work in grammaticalisation may have focused on studying the development of individual items, the end of the 20th Century saw an increasing interest in the role of the *construction*. Concomitantly, there has been a noticeable increase in studies dealing with questions of language change from a Construction Grammar perspective (see, e.g., the papers in Bergs & Diewald 2008, Trousdale & Gisborne 2008, and Barðdal et al. 2015, as well as Fried 2013, Barðdal 2013, De Smet 2013, Hilpert 2013). This section will first discuss some relatively recent changes in the framework of Construction Grammar. We will then expand on the notion of productivity, which has come into focus as an important aspect to take into account when describing the behaviour and development of a construction. The topic of productivity naturally brings us to the issue of the constructional network. In the third paragraph, it will be shown that a construction can be conceptualised as a taxonomic network of schemas and subschemas which exhibit varying degrees of productivity at different levels of abstraction. Crucially for current purposes, the degree of productivity of each of these subschemas is subject to diachronic shifts, which may cause a reorganisation of the internal structure of the network.

2.1.1 Constructions in variation and change

Construction Grammar is an umbrella term that covers a variety of cognitively informed, mostly (though not exclusively) usage-based approaches that accord a central role to the notion "construction", which is defined as a conventionalised form-meaning pairing (see, e.g., Fillmore et al. 1988, Goldberg 1995, Croft 2001, Langacker 2005). Although the different constructionist approaches differ in important respects, such as the way in which they formalise constructions (see the *Oxford Handbook of Construction Grammar* edited by Hoffman and Trousdale 2013 for an overview of the different branches in Construction Grammar), there are some main tenets that are shared by all constructionist strands (Goldberg 2003, 2013). The most fundamental of these is that language is argued to be a repository of constructions, i.e. form-meaning pairings, that are organised in a taxonomic network that has been termed the "constructicon" (Jurafsky 1992). These constructions may vary in terms of their internal complexity and in terms of their schematicity, but there is no qualitative difference between lexical constructions on the one hand and procedural or grammatical constructions on the other – that is to say, Construction Grammar does not assume a sharp distinction between lexicon and grammar. The formal complexity of a construction is defined in terms of a continuum

that ranges from atomic, e.g. morphemes and monomorphemic words, to complex, multi-word grammatical patterns. The schematicity of a construction is described in terms of its phonological specificity, which is likewise a matter of degree. On one end, there are substantive constructions, which are fully specified, e.g. the lexical items *cat* or *house*, on the other end, we have fully schematic constructions. A prime example of the latter are the so-called argument structure constructions, such as the ditransitive or double object construction [SUBJ V OBJ_1 OBJ_2], but we also find the syntactic categories [NOUN] or [VERB] at this end of the continuum. In between, we find a wide range of partially schematic constructions that contain both substantive slots and "empty", schematic slots, e.g. the 'What is X doing Y?'-construction (Kay & Fillmore 1999).

In the early days, much work in Construction Grammar was focused on "unusual" or "idiosyncratic" constructions, which, in Chomskyan theories, had been relegated to the periphery of grammar. This is obvious from Goldberg's (1995: 4) early definition of a construction: "Phrasal patterns are considered constructions if something about their form or meaning is not strictly predictable from the properties of their component parts or from other constructions." About a decade later, however, this definition has been expanded to include fully compositional patterns as well: "In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency" (Goldberg 2006: 5). Construction Grammar, as a theory of language, clearly intends to account for all possible expressions in the grammar of any language. Until fairly recently, however, Construction Grammar tended to focus mainly on present-day language; aside from some superficial remarks and early exceptions (e.g. the paper by Israel 1996 on the diachronic development of the *way*-construction in English), issues of lectal variation or language change were not addressed. Since the mid-2000s, constructionist approaches have widened their scope to include the study of variation and change in constructions: there has been a marked increase in studies tackling the emergence of new constructions and subsequent (long-term) constructional changes, as well as a rising interest in non-standard language varieties (see, e.g., Siewierska & Hollmann 2007, Barðdal 2007, Hoffmann & Mukherjee 2007, Grondelaers et al. 2008 and the papers in part 5 of Hoffman & Trousdale 2013 and in Boogaart et al. 2014). In the next two subparagraphs, we will first explore the study of regional or national variation in constructions and then discuss the emergence of Diachronic Construction Grammar as a "new" approach to language change.

2.1.1.1 Variation: Cognitive Sociolinguistics and Construction Grammar

An important tendency that has contributed to the increased interest in the (synchronic) variational dimensions of constructions is the introduction of Cognitive Sociolinguistics, a research framework which is aimed at the empirical investigation of the socio-cultural dimensions of linguistic variation by combining the methodologies and objects of study

from sociolinguistics and Cognitive Linguistics into one integrated framework (Boogaart et al. 2014: 6). It brings together the basic idea of Cognitive Linguistics that the structure of language is determined by underlying cognitive mechanisms (Croft & Cruse 2004), and the practices of variationist sociolinguistics, which is primarily concerned with language as a social structure, i.e. with how social factors shape the language system (see the papers in Kristiansen & Dirven 2008, and Geeraerts et al. 2010, as well as, e.g., Harder 2010, Pütz et al. 2012). As a usage-based approach to language, Cognitive Sociolinguistics seeks to develop empirical tools for investigating real language in use, i.e. language as a product of language users interacting in specific communicative situations. One corollary of this new research framework is that there has been a noticeable increase of studies dealing with intralingual variation in constructions, that is, variation in the use of certain constructions in different varieties of one and the same language. Given that English is one of the most widespread languages across the world, it is no surprise that differences between regional and national varieties of English are especially well-studied. There are a number of studies that go beyond the known lexical differences between American and British English (e.g. American *subway* vs. British *underground*) and focus on more structural differences. Mukherjee & Hoffmann (2006) (and later Hoffmann & Mukherjee 2007) have investigated how Outer-Circle varieties of English, so-called New Englishes, deviate from British English with respect to verb complementation. Focusing on ditransitive patterns in Indian English in particular, they find that there are certain differences with respect to the complementation of some typical ditransitive verbs like *give* and *send*, and with respect to the variety of verbs found in ditransitive patterns. For example, the verbs *give* and *send* appear to be less prototypical ditransitive verbs in Indian English because they are more often found in a pattern with just one explicit object, whereas the pattern with two explicit objects (i.e. the basic ditransitive pattern) is by far the most frequent one in British English. In Indian English, then, the basic ditransitive pattern is instantiated by a range of verbs that do *not* take the ditransitive construction in British English e.g. *provide*, *supply*, *present*, *rob*, *notify*, etc. which usually occur with a prepositional complement with *with* or *of*. It is not unusual that varieties of the same language display diverging preferences in terms of the verbs that are used in specific constructional patterns. For the *into*-causative, Wulff et al. (2007) use a distinctive collexeme analysis to find out whether the top verbs that are associated with the construction are different in American and British English. The analysis reveals a subtle but interesting pattern: the *into*-causative is primarily associated with verbs of verbal persuasion in American English, whereas British English shows a stronger preference for physical force verbs and other negatively connoted verbs, e.g. verbs of negative emotions. Szmrecsanyi (2010) compares the use of the *s*-genitive and the *of*-genitive in American and British English by examining which factors determine the choice of genitive in both varieties. With respect to the overall use of genitives, he observes that the *s*-genitive is more frequent in American English than in British English in journalistic data, but the

situation is reverse in spoken data. These frequency differences, he argues, are epiphenomenal to underlying language-internal mechanisms that condition the genitive choice. Although the conditioning factors (e.g. animacy of possessor, length of possessor, etc.) are largely the same across text types and geographical varieties, the magnitude of the effect can show substantial variation if language-external factors are factored in.

Due to the heavy focus on varieties of English, there is as of yet not enough empirical evidence to ascertain whether intralingual variation is cross-linguistically pervasive. Still, we do expect to find similar national variation in Dutch, especially given the very different standardisation histories of Belgian and Netherlandic Dutch. Without going into too much detail here, Netherlandic Dutch has gone through a process of regular standardisation, which started in the 16th Century and was more or less completed by the beginning of the 20th Century, whereas Belgian Dutch is characterised by a delayed standardisation. When the standardisation process started in the northern part of the language area (viz. the Netherlands) in the 16th Century, the southern provinces (viz. Belgium) remained under Spanish control and could not participate in the standardisation. Dialects of Dutch were still spoken by a large part of the population, but the official language was French. As Belgium subsequently came under Austrian and then French rule, social and political factors continued to prevent the standardisation of Belgian Dutch until the 19th Century. In the 19th Century, the Flemish Movement began to defend the rights of Dutch in Belgium, but there was disagreement between those who strove for a separate "Belgian" Dutch standard language and those who advocated the adoption of the already available standard language of the Netherlands. Eventually, Belgium officially introduced the (Netherlandic) Dutch standard as its national standard, in addition to the official languages French and German (see De Vooys 1975, Van den Toorn 1997, Van der Horst & Marschall 2000 for a comprehensive account of the history of the Dutch language). Despite the shared "official" standard, there continue to be differences between the national varieties of Dutch. Similar to the situation of American and British English, most speakers of Belgian or Netherlandic Dutch are aware that there is considerable variation between both national varieties with respect to pronunciation and lexicon (see Debrabandere 2005 for an overview of all types of lexical differences, or for more specific studies, see Geeraerts et al. 1999 on football and fashion vocabulary, or Speelman et al. 2008 on varying adjective choices). Perhaps less noticeable to the everyday language user, but all the more so to the linguist, there are also (sometimes subtle) differences in the usage patterns of syntactic constructions in Belgian and Netherlandic Dutch. Often, the external factor of national (or regional) variety primarily receives attention insofar as it interacts with language-internal factors, for example in governing the choice between two functionally (near-)equivalent constructions (cf. the genitive construction in Szmrecsanyi 2010). In their study on *er*-insertion ('there'), Grondelaers et al. (2008), for example, demonstrate that there are differences between the national varieties of Dutch that cannot be solely explained by underlying cognitive

structures or language-internal factors. Their findings suggest that external factors such as national variety or register may have an effect on (syntactic) variation in different ways. First of all, they find that external factors may interact with language-internal factors: the presence of a locative adjunct (which is a language-internal factor) is a stronger cue for *er*-omission in Netherlandic Dutch than in Belgian Dutch. Second, external factors can also have a more direct or independent effect in that one variant is used more often in one variety than in another, even if the effects of other parameters are controlled for. In the case of *er*-insertion, Grondelaers et al. note that *er* 'there' is used more often in Belgian informal registers, whereas there is no clear independent effect of register in Netherlandic Dutch. The role of external factors like national variety and register is also highlighted in Speelman & Geeraerts's (2009) study on the distribution of the causative verbs *doen* 'to do' and *laten* 'to let' in spoken varieties of Belgian and Netherlandic Dutch (e.g. *De politie liet/deed the auto stoppen* 'The police made the car stop'). Based on previous studies, they formulate the hypothesis that *doen* is more frequently used in Belgian Dutch than in Netherlandic Dutch. Within Belgian Dutch, then, it is expected to be even more frequent in informal sources because typically Belgian forms have been found to be used more often in informal registers (Speelman & Geeraerts 2009: 175, based on Geeraerts et al. 1999, also see Grondelaers et al. 2008). Their results do confirm the first part of the hypothesis: as expected, the older form *doen* 'to do' is more frequently used in Belgian Dutch, which is considered to be the more archaic variety due to its delayed standardisation (the higher frequency of *doen* in Belgian Dutch is later reconfirmed for journalistic data as well, see Levshina et al. 2014). However, unlike in the case of *er*-insertion, the Belgian informal data do not contain a particularly higher proportion of *doen* than other genres, so the second part of the hypothesis is not confirmed. On the contrary, *doen* being the older form, it is actually associated with formal, written genres and both Belgian and Netherlandic Dutch seem to disfavour its use in spontaneous conversation. Colleman (2010) also accords a central position to national variation by focusing on the different semantic potential of one specific construction, viz. the benefactive ditransitive, in Belgian and Netherlandic Dutch (e.g. *Hij schenkt zich een glas in* 'He pours himself a drink', *Moeder schepte ons een bord groenten op* 'Mother dished us up a plate of vegetables'). He proposes that, in view of the constructionist tenet that there is no fundamental difference between lexical items and argument structure constructions (cf. supra), it need not surprise us that constructional meanings can display the same kind of lectal variation that is found in concrete lexemes. Previous studies had already reported that there was regional variation in the use of the benefactive ditransitive construction in that southern (and some eastern) dialects and regional substandard varieties allow for a wider range of uses than Standard Dutch (Colleman & De Vogelaer 2003), but it had not yet been investigated to which extent this wider distribution also applies to the standard variety of Dutch spoken in Belgium compared to standard Netherlandic Dutch. Though the number of benefactives retrieved from the corpus is

small, the data do confirm that the benefactive is used with a more diverse set of verbs in Belgian Dutch. As a possible explanation, Colleman (2010: 206) proposes that in Netherlandic Dutch, the act of preparing the transfer and the transfer itself need to be almost coincidental – much like with regular verbs of giving – to get an acceptable benefactive, whereas this constraint is less strict in Belgian Dutch and even less so in the dialects. This may explain why a sentence like *Grootmoeder heeft me een trui gebreid* 'Grandmother has knitted me a sweater', in which the act of knitting and the act of giving are not coincidental, is acceptable in Belgian Dutch but not (anymore) in Netherlandic Dutch. Other language-internal or external factors, such as the concreteness of the transfer or register variation, may also come into play, but their exact role still requires further research. Studies such as the above show that it is worth exploring national variation in the Dutch intensifying fake reflexive resultative construction as well. In addition, there has also been some interest in the use of intensifiers in Belgian and Netherlandic Dutch, to which we will return in §3.3.1.

2.1.1.2 Change: from Construction Grammar to Diachronic Construction Grammar

The adoption of concepts, viewpoints and terminology from constructionist theories of grammar to the study of language change has come to be called "Diachronic Construction Grammar". Given that the study of language change was mainly couched in terms of grammaticalisation at that time, the emergence of this "new" framework inevitably prompted some discussion on the relationship between the grammaticalisation framework and Diachronic Construction Grammar (see Noël 2007, Traugott 2008a, b, Trousdale 2010, 2014, Fried 2013, Hilpert 2015b, and the papers in Coussé et al. 2018 for further discussion). It must be stressed that both approaches share important common ground, and many of the concepts that were central to grammaticalisation theory – such as analogy, reanalysis, gradience and gradualness – remain relevant in Diachronic Construction Grammar.The fact that grammaticalisation has recently incorporated some basic constructionist ideas has strongly intertwined both diachronic theories. This "rapprochement" between constructional approaches and grammaticalisation has been described as one of the important trends that have contributed to the increase in research scope and which have brought a fresh wind to construction grammatical research (Boogaart et al. 2014: 6-7, Colleman & Van de Velde 2015: 137-138). In general, the discussion concludes that grammaticalisation – in its traditional sense of the process through which "lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions" (Hopper & Traugott 2003: 1) – is just one of several processes of change that constructions may undergo. Diachronic Construction Grammar offers a unified framework that can account for a wider range of diachronic shifts than the kinds of

phenomena that have typically been the focus of grammaticalisation research. These diachronic shifts are taken together by Hilpert (2013) under the encompassing term "constructional change" – not to be confused with the slightly different notion of constructional change in Traugott & Trousdale (2013) (cf. infra). An important addition to the research field are lexical semantic developments, lexicalisation processes or other changes within the field of lexis which are largely side-lined in the grammaticalisation framework (see Lehmann 2002, Himmelmann 2004, Trousdale 2008a, b for some discussion on the relation between grammaticalisation and lexicalisation). While grammaticalisation research has always been chiefly concerned with grammatical constructions, the question of lexical versus grammatical constructions is less crucial in Diachronic Construction Grammar: any change to the formal and/or functional aspects of a construction falls under the purview of constructional change (Noël 2007, Hilpert 2013). Within the domain of syntactic changes as well, Diachronic Construction Grammar offers a broader perspective than traditional grammaticalisation theory in that it may also include the study of construction-specific word order changes, derivational morphology and different types of frequency changes, among others (see, e.g., Booij 2010a on constructional morphology, and Hilpert 2013, Diessel & Hilpert 2016, Hilpert 2017 on frequency developments). Recently, there has been some discussion with respect to the research object of Diachronic Construction Grammar and the aims that Diachronic Construction Grammar aims (or should aim) to achieve (Hilpert 2018). In general, Diachronic Construction Grammar has been much less explicit about its cognitive commitment than (cognitive) Construction Grammar. In most of the studies that will be discussed below, Diachronic Construction Grammar is treated as a descriptive framework for investigating all kinds of language changes, but the cognitive implications of these changes are often left implicit.

One of the fundamental works in Diachronic Construction Grammar is the monograph by Traugott & Trousdale (2013), in which they posit a distinction between constructionalisation and constructional changes. Constructionalisation is defined as the formation, typically through a succession of neo-analyses, of a new type node in the constructional network, i.e. of a new conventionalised form-meaning pairing (2013: 22). Before it reaches the status of a full-fledged construction, a constructionalising element is argued to undergo several changes called "pre-constructionalisation constructional changes". These generally involve a number of small local changes in context that affect the formal or semantic features of the constructionalising element: a formal pattern can, for instance, receive multiple interpretations, depending on the context it is used in. This step bears a striking resemblance to the "critical contexts" or "bridging contexts", proposed by Diewald (2002) and Heine (2002) respectively, in the grammaticalisation framework. This new interpretation may be arrived at via a process of pragmatic inference at first, but as it gains salience, it may become fully semanticised: the formal dimension has remained unchanged, but a new meaning has arisen. In order to resolve

this form-meaning mismatch, the formal dimension needs to be aligned with the semantics and a new construction is created. Post-constructionalisation constructional changes, in their turn, often manifest themselves as morphological or phonetic reduction and a relaxation of the constraints that hold over the different slots in the construction, leading to type expansion – cf. the concept of host-class expansion in Himmelmann (2004).

Much of the existing work in Diachronic Construction Grammar is concerned with how new constructions come into being, or, in other words, how certain patterns "constructionalise" into new form-meaning pairings. Of course, the development of a construction does not stop at the moment of constructionalisation: established constructions continue to undergo both formal and semantic changes that are worthy of linguistic attention. Although the emergence and early stages of development of new constructions remain an interesting topic of research, researchers in Diachronic Construction Grammar have begun to turn their attention to a wider range of diachronic changes, bringing us back to Hilpert's broader notion of constructional change.

> Constructional change selectively seizes a conventionalized form-meaning pair of a language, altering it in terms of its form, its function, any aspect of its frequency, its distribution in the linguistic community, or any combination of these. (2013: 16)

Colleman & De Clerck (2011) focus on semantic shifts in the double object construction [DOC], by comparing the semantic range of application of the DOC in 18th-Century English to its present-day semantics. It appears that the construction has been subject to a process of semantic specialisation, in that a number of uses/meanings have disappeared since the early stages of Late Modern English. The obsolescence of certain subsenses is attested for equivalents of the DOC in other Germanic languages as well (see, e.g., Colleman 2010 on the Dutch benefactive, Barðdal et al. 2011 on the Mainland Scandinavian languages). A construction can also gain new subsenses over time, as is demonstrated by the case of the *para*-Infinitive in Brazilian Portuguese (Torrent 2015). Aside from cases of semantic specialisation or generalisation, subsenses may come to occupy a more or less central position in the semantic network without affecting the array of constructional subsenses. Geleyn (2016) shows that the semantic range of the Dutch *aan*-construction has not undergone important qualitative changes since the early stages of Modern Dutch, but there are subtle quantitative shifts with respect to the frequency of use of some subsenses. Given the long history of a construction like the DOC, which already existed in Old English (De Cuypere 2014), or the *para*-Infinitive construction, which dates back to the 13th Century, it is hardly surprising that the semantics and the distributions of these constructions in present-day English and Portuguese are remarkably different from those observed in older stages of the respective languages. However, it may be worthwhile to look at what happened after the initial constructionalisation of constructions that have entered the grammar at a more *recent*

time as well. Colleman (2015), for example, looks into the mid- to late 20th Century semantic changes and expansion of the so-called *krijgen*-passive, which was only introduced in the late 19th Century. In its rather short period of existence, the *krijgen*-passive has already undergone a number of semantic evolutions. In section 2.1.3 below, taking a constructional network perspective, we will return to some of these examples and argue that semantic changes that are often described in terms of loss or gain of certain subsenses actually involve the emergence or obsolescence of certain subschemas within the network. Before we get to that, we want to have a closer look at another topic that has come to the fore as being of particular interest in Diachronic Construction Grammar, viz. (syntactic) productivity. The next section will discuss how the term productivity has been used in previous linguistic research and how our use of productivity, as applied to constructions, fits within this wider concept.

2.1.2 Productivity

One of the most elaborate overviews of the notion of productivity in linguistic research is presented in Barðdal (2008), who kicks off her monograph on productivity with the observation that "productivity has presented one of the most indistinct and puzzling phenomena for linguistic research over the decades" (2008: 9). Based on the different uses and senses of productivity found in the linguistic literature, Barðdal discerns three different concepts of productivity, that need to be disentangled so as to avoid terminological or conceptual confusion. The first concept concerns productivity as *generality*, which is further specified as "having a wide coverage", "schematic" or "default". Second, the productivity as *regularity* concept encompasses senses like "rule-based", "transparent" or "operative". In our investigation, productivity will be thought of in terms of the third concept, *extensibility*, i.e. the extent to which a construction is able to attract or expand to new items. As a large share of the studies on linguistic productivity as extensibility, as well as some important empirical measures are situated within the domain of morphology, we start with a brief survey of morphological productivity before turning to constructional productivity. We close this section with a discussion of the diachronic implications of productivity and the way in which productivity will be implicated in this thesis.

2.1.2.1 Morphological productivity

Morphological productivity has been defined as follows:

> Onder produktiviteit als morfologisch fenomeen verstaan we dan de voor taalgebruikers bestaande mogelijkheid door middel van het morfologisch procédé dat aan de vorm-betekeniscorrespondentie van sommige hun bekende woorden ten

> grondslag ligt, onopzettelijk een in principe niet telbaar aantal nieuwe formaties te vormen. (Schultink 1961: 113)
> 'Productivity as a morphological phenomenon is understood as the possibility for language users to unintentionally coin new words by using an existing morphological pattern (that is, a word-formation rule that lies behind the form-meaning correspondences of a number of words they are already familiar with). The number of new formations that emerges as a result is in principle uncountable.' [my translation, EG]

Although there is general consensus on the idea that a word-formation pattern is to be considered productive if it can be extended to new words that enter the language, the criteria of unintentionality and uncountability have provoked some discussion. The unintentionality criterion implies that any "deliberate", creative neologisms that are created for humoristic or playful purposes should be taken out of the equation when discussing the productivity of a word-formation pattern (Schultink 1961, Lieber 1992). However, the notion of unintentionality is rather subjective and it is not always easy to decide which words were created deliberately and are perceived as such, and which words simply pass unnoticed (Plag 1999, Dal 2003, Booij 2009). The criterion of uncountability is also problematic because it is unlikely that a morphological pattern could really produce an *infinite* amount of new formations: there are always at least some restrictions that constrain the applicability of a pattern, even if the pattern is unmistakably "extensible" to new words and, therefore, productive. The finding that a pattern can be *more or less* productive casts doubt on the crisp categorical distinction between productive versus unproductive rules, and has led to the introduction of the notion of gradience and "gradability" in the study of morphological productivity (Baayen 2009). Even so, there were no clear criteria in place to measure these different degrees of productivity, and early studies were mostly based on the researcher's own introspection. Such introspective judgments proved to be unreliable, as researchers often had diverging intuitions on the productivity of one and the same word-formation pattern. For instance, whereas Schultink (1962) observes that the nominal suffix *-te* (more or less equivalent to the English suffix *-ness*) has ceased to be productive in Dutch, Booij (1977) finds that it is still sporadically attached to adjectives to create new nouns, which suggests at least some degree of productivity. This has prompted morphologists to search for empirical, corpus-based measures that can deal with productivity as a gradual phenomenon and lend some objective support to those linguistic intuitions. In this section, we restrict ourselves to a brief introduction of the methods that are relevant to the present investigation. In Chapter 4, §4.3, these measures will be discussed in a more detailed fashion.

The first attempt at an objective measure of productivity is found in Aronoff (1976), who takes the ratio of all words that are actually formed by a given word-formation rule and all words that could potentially be produced as the output of that rule. This ratio is formalised by Baayen (1989) as I=V/S (I stands for Index of Productivity, V for

types/actual words and S for potential words). One of the limitations of Aronoff's suggested method of measuring productivity lies in the rather vague definition of the "actual" and "potential" words. It is not clear whether the actual word count pertains to a speaker's mental lexicon, a list of words in some dictionary or in a fixed corpus. The potential output of a word-formation pattern, Aronoff argues, is based on the number of possible bases for the pattern, but he does not explain how to delimit this set of possible bases (Baayen & Lieber 1991: 804). Aronoff (1983) later tries a different approach and suggests that productivity can be measured in terms of a negative correlation with the token frequencies, i.e. the total frequency of occurrence (in a corpus), of the output lexemes of different word-formation rules. A word-formation rule that produces items with a lower token frequency is supposed to be more productive than a rule that generates high frequency words. The idea that token frequency detracts from productivity continues to hold some sway: it is often argued that lexemes with a high text frequency are not produced on the fly and are analysed as a whole, which implies that they are unlikely to be the result of a productive pattern (Bybee 1985, 1995, Bybee & Thompson 1997, Clausner & Croft 1997). In contrast to token frequency, type frequency, i.e. the number of *distinct* items found in a corpus, has been widely accepted as a good indicator of productivity: if a pattern is instantiated by many types, the representation of the pattern itself grows stronger and, consequently, it is more likely to be extended to even more types (Bybee 1985, 1995, Bybee & Thompson 1997). Still, there may be a way in which both high type and high token frequencies contribute to productivity – at least in a constructional model – but we will come back to this when we are discussing Barðdal's (2008) model of productivity below. Perhaps the most influential measures have been introduced by Harald Baayen, who wrote his dissertation (Baayen 1989), as well as several journal articles on a corpus-based approach to morphological productivity (see also Baayen 1990, Baayen & Lieber 1991, Baayen 1992, 1993, 2009, inter alia). By adopting a multidimensional view on productivity in his body of work, Baayen may very well have been the first to fully acknowledge the complexity of the notion. He sets himself apart from previous accounts by introducing a distinction between two different, but equally important aspects of productivity, viz. realised productivity and potential productivity. The realised productivity is captured by the type frequency, which gives us an idea of the current extent of use of the pattern under investigation. However, it does not make any predictions about the potential (or future) expansion to new types. To give but one example, Bürgisser (1983) observes that despite the abundance of nouns ending in *-nis* in present-day German (e.g. *Erlebnis* 'experience', *Geheimnis* 'secret', *Gefängnis* 'prison', *Ärgernis* 'annoyance'...), this suffix is no longer used to coin new formations. The high type frequency suggests that the word-formation pattern with *-nis* must have spawned a variety of new deverbal nouns at some point, but fails to give any information about its availability to *still* do so. In the same vein, a new or recent pattern may have a low type frequency at present, but it could have a high chance of attracting many more types. The

extensibility of the pattern is estimated by the potential productivity measure ($\mathscr{P}$), the ratio of hapax legomena [HL], i.e. the types that occur only once in the corpus, created by a word-formation rule to the total number of tokens created by that rule. This measure is based on the idea that a high number of hapax legomena and a low number of higher-frequency tokens positively influence the extensibility to new types. Both measures, each highlighting a different aspect of the productivity of a morphological pattern, are brought together in a model that allows us to say something about the global productivity of a pattern. This global productivity is represented as a graph in which the potential productivity is plotted on the horizontal axis and the realised productivity is plotted on the vertical axis (Baayen & Lieber 1991, Baayen 1992).

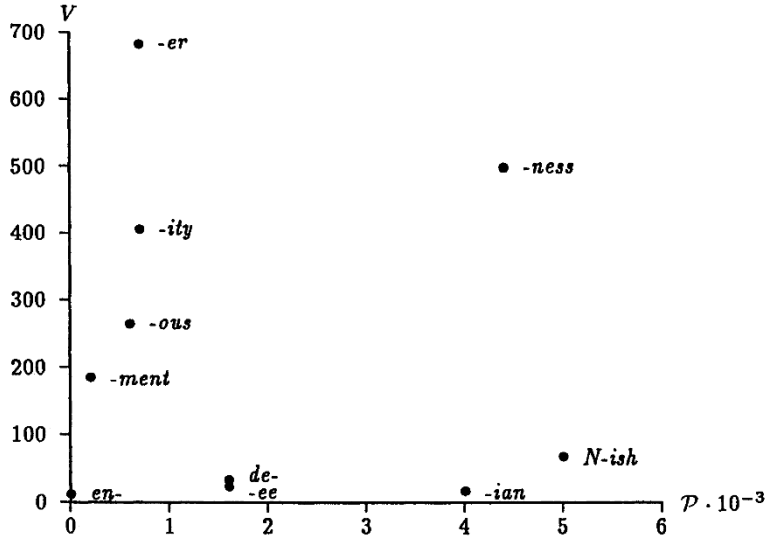


Figure 2.1. Global productivity graph of a number of English word-formation processes as found in the Cobuild corpus (adopted from Baayen 1992: 124)

The visual representation in Figure 2.1 makes it is possible to quickly gauge the differences in productivity between different rules: the (globally) more productive rule is situated more to the top right with both high $\mathscr{P}$- and V-values while the less productive rule, which has lower $\mathscr{P}$- and V-values, will be situated to the bottom left of the plane. In what follows, we move on from morphological productivity to the current issue at hand, viz. constructional productivity.

2.1.2.2 Constructional productivity

The first mention of productivity that is not strictly limited to morphology is found in Hockett (1958: 307): "The productivity of any pattern – derivational, inflectional, or syntactical – is the relative freedom with which speakers coin new grammatical forms by it." Some forty years later, Langacker (1999: 114) gives his own definition of productivity that can be applied to syntactic structure (although he does not refer to "syntactic productivity" in those words): "Productivity is a matter of how available a pattern is for

the sanction of novel expressions." Still, we have to wait until Barðdal (2008: 1) for a first real definition of productivity that is explicitly applied to (syntactic) constructions:

> By syntactic productivity, I refer to an argument structure construction's ability to attract new or existing lexical items, i.e. a construction's extensibility.

The clear parallel between the definitions of morphological and syntactic productivity lies in the emphasis on the extensibility of an existing pattern to create new formations. However, the definition of constructional productivity does not mention the unintentionality of the process, an aspect that was crucial in Schultink's (1961) definition and in morphological productivity overall. In the morphological tradition, Bauer (2001: 65) clearly distinguishes productivity from creativity, defining creativity as "non-productive innovation", and he reserves the term analogy, rather than productivity, for the extension of such non-productive patterns. This dichotomy between productivity and creativity, or productivity and analogy, is rejected in constructionist approaches, which support a gradient approach to productivity, rather than viewing productivity as an all-or-nothing phenomenon. Barðdal (2008: 3) suggests that analogy and productivity are "two sides of the same coin" and that all types of extensions should be taken into account. Zeschel (2012: 228) also argues in favour of the inclusion of creative, unconventional uses of a construction in the study of syntactic productivity because they can tell us something about the language user's extension strategies and the existing linguistic conventions. In contrast, Zeldes (2012: 40-43) suggests that we should consider teasing apart productivity from creativity in syntactic research just like in morphological research, although he offers no straightforward alternative to what he calls the "impractical criterion" of speaker intentionality or awareness. Following Langacker's (1987) definition of linguistic creativity below, we are inclined to claim that creativity, in the sense of speakers "toying around" with existing constructions, may actually be a symptom of productivity.

> The creation of novel expressions, including extensions (involving figurative language, the adaptation of linguistic units to new situations*, or even willful violation of convention*) and also the straightforward computation of fully sanctioned expressions. (Langacker 1987: 490; emphasis added)

These novel and creative uses of a constructional schema corroborate the idea that the constructional schema can contribute its own, independent constructional meaning regardless of the lexical items that fill the open slots. A case in point is the *(all)* [[X]-*ed*]$_V$ *out* construction (described by Jackendoff 2013), which Audring & Booij (2016: 7) mention as an example of type coercion in constructions. Examples like *By midnight:30 I was all Amsterdammed out* or *Just in case you're not all Biebered out already, here's the full studio version of "Mistletoe"* show that the abstract schema has become so productive that the semantics of 'being exhausted from X-ing to excess' have become predictable, irrespective of the inserted item. It must be stressed that the word-formation patterns that are studied in

the field of morphology are not all that different from the constructions that are at the centre of research in Construction Grammar. As Construction Grammar has abandoned the strict division between lexicon and grammar, morphological rules can readily be reinterpreted as constructional schemas that contain information about the shared form and meaning of sets of complex word forms (see the work by Geert Booij on Construction Morphology, e.g. Booij 2010a, b). For instance, the word-formation rule that derives nouns from adjectives by adding the suffix *-ness* can be rendered as a constructional schema $[[X]_A$ *ness*$]_N$ ⇔ 'the property/state of A'.[1] This schema has arisen as a generalisation over existing *-ness* nouns, and it is considered to be productive because new nouns may be formed if the X-slot is filled by a novel, concrete adjective. The concept of empty slots that are instantiated by concrete lexical material in real language use carries over to multi-word expressions and argument structure constructions. Argument structure constructions can be partly lexically specified, as is the case for the *way*-construction [SUBJ$_i$ [V [POSS$_i$ *way*] OBL] ⇔ 'create and/or follow a path by means of V/while V'ing', or they can be fully schematic constructions, containing only empty slots, such as the double object construction [SUBJ V OBJ$_1$ OBJ$_2$] ⇔ 'cause to receive by means of V'. In a situation in which we are confronted with multiple open slots, the extensibility of some slots may be more relevant to productivity than others. When Barðdal (2008: 29) proposes that the productivity of an argument structure construction refers to its extensibility to new types, she explicitly adds that a type, in that context, should be regarded as a verb or a predicate. In other words, a construction is considered to be productive if it can attract novel verbs. In argument structure constructions that are "verbal" in nature (e.g. the ditransitive construction, causative constructions or motion constructions in which the verb is the central element), the productivity of the entire construction does appear to hinge upon the extensibility of the verb slot. However, depending on the research questions and the construction under focus, a different, non-verbal slot may be of equal or even greater interest (Zeldes 2012). In addition, it can be meaningful to consider the productivity of multiple slots at different hierarchic levels. That is, after having determined the verbs that are compatible with the verb slot of a certain construction, for example the DOC, at the highest level of abstraction, one could drop down to a lower level of abstraction to compare the productivity of the OBJ$_2$-slot in combination with those different verbs or verb classes (see the next section on the hierarchic organisation of the constructional network). There is no unequivocal answer to the question of which slot one should focus on first; it is up to the researcher to select the slot(s) that is (are) the most appropriate for their research purposes.

[1] The ⇔ symbol is the conventional way of capturing the association between the form (on the left of the arrow) and the meaning (on the right) of a construction.

The question now arises as to how we can empirically measure this constructional productivity. As there are no fundamental differences between the way in which the empty slots in morphological patterns and argument structure constructions are filled, it does not come as a surprise that the same productivity measures that are used to compute morphological productivity can be fruitfully extended to the instantiation of constructional slots in argument structure constructions. In his book-length study of syntactic productivity, Zeldes (2012: 106-125) discusses a number of case studies in which he applies Baayen's measures (cf. supra) to concrete syntactic phenomena. A first case study involves the use of the construction with *wegen* + N ⇔ 'because of N' in German, which can be used either prepositionally or postpositionally without difference in meaning (e.g. *wegen des Vaters* vs. *des Vaters wegen* 'because of the father'). Given that the (older) postpositional use has receded to formal registers and a number of formulaic expressions in present-day German, it is felt to be less productive than the prepositional variant. The frequency-based productivity measures that were briefly introduced earlier all confirm the intuition that the postpositional construction is rarer than the prepositional construction and less likely to be extended to new contexts. Another case study is concerned with the direct object selection of a set of English transitive verbs, viz. *drink, eat, spend, incur, sift*, etc. A remarkable finding is that the results from the different productivity measures now lead to very different productivity rankings. Although frequent verbs like *drink* or *eat* show up with a larger argument spectrum – i.e. they have a higher degree of realised productivity – than an infrequent verb like *sift*, *sift* actually has a higher chance of being encountered with a previously unattested argument in the corpus, i.e. a higher degree of potential productivity. The conclusion to be drawn from these results is that syntactic productivity, just like morphological productivity, is too complex to be reduced to a single measure; a multidimensional approach to productivity that includes multiple complementary measures makes it possible to highlight different dimensions of productivity.

The examples above have demonstrated that the frequency-based measures by Baayen and colleagues already provide some empirical foundation to the study of linguistic productivity. However, within a constructionist model of productivity, Baayen's productivity complex is "missing" or ignoring at least one important factor, viz. the qualitative or semantic aspect of productivity. Goldberg (1995), for example, attempts to explain the limits with respect to the distribution of constructions by positing semantic constraints. One example of such a constraint on the resultative construction is that "the resultative adjective must denote the endpoint of a scale" (1995: 193). The following examples do not meet this constraint and are therefore marked by Goldberg as ungrammatical: **He drank himself happy*, **He wiped it damp*, **The bear growled us afraid* or **He hammered the metal beautiful*. This constraint explains why the range of adjectives that can occur as a resultative phrase are so limited – or, in other words, why the slot of the resultative phrase is not very productive. Boas (2003, 2005) argues that Goldberg's

semantic constraints are too general to accurately predict the distribution of the resultative construction. In response, he introduces a set of more fine-grained semantic, pragmatic and syntactic restrictions that are conflated in so-called "mini-constructions" for each verb that can be used in the resultative construction. These mini-constructions inherit formal and functional properties from higher-level schemas, but they add extremely specific information for each verb sense in the form of an event-frame (Boas 2003, 2008). Furthermore, collocational restrictions (i.e. restrictions on the productivity of a constructional slot) can be influenced by the world knowledge that is associated with certain verbs and concepts (Goldberg & Jackendoff 2004, Boas 2005).[2] Zeldes (2012) also finds that lexical semantics and world knowledge can affect the range of lexemes we expect to see in certain positions: for example, in as far as there are more "edible" things than "bakeable" things, the verb *to eat* should be more productive than the verb *to bake*. However, if productivity was primarily determined by such semantic and pragmatic constraints, we would expect functionally and semantically related constructions to exhibit similar degrees of productivity. As this is often not the case, Zeldes concludes that productivity cannot be predicted on the basis of such constraints alone.

In her monograph on the productivity of Icelandic case constructions, Barðdal (2008) seeks to combine structural and semantic aspects in a new model of productivity. The alternative she proposes factors in the type frequency and the semantic coherence of a schema, which, taken together, should accurately predict a schema's productivity. Type frequency is now defined as the number of types which can fill a specific slot in a construction; schema coherence concerns the internal consistency between the members/types that can fill that slot. With respect to argument structure constructions, this coherence is mainly semantically defined, although there exist some cases of morphophonetic restrictions (Barðdal 2008: 27). The idea that type frequency positively correlates with productivity has been repeatedly stated in the literature (Bybee 1995, Goldberg 1995, Bybee & Thompson 1997, Clausner & Croft 1997, among others), but we have already seen that type frequency alone is only part of the story. Even Bybee (1995), who makes a strong case for the importance of type frequency, admits that this is not the only factor that needs to be taken into account: "a pattern cannot attain full productivity if there are restrictions – phonological, semantic or morphological – on its applicability" (1995: 435). The role of semantic coherence in productivity is not entirely new either (it is mentioned in passing in, e.g., Aronoff 1976 and Baayen 1992), but the novelty of Barðdal's approach lies in the specific interaction between type frequency and semantic

[2] At the same time, world knowledge or contextual background knowledge does not only constrain the applicability of a construction, it can also explain why sentences which are generally unacceptable instances of a pattern, can sound perfectly fine given the right context. This is what Boas (2011) calls "leaking": in certain contexts, the conventionalised argument structure specifications of a verb can *leak* so that "unacceptable" non-conventionalised utterances (e.g. *Ed hammered the metal safe*) are judged acceptable.

coherence. Barðdal (2008) hypothesises that there is an inverse correlation between the type frequency of a construction and its degree of semantic coherence. This amounts to the idea that the importance of semantic coherence for the productivity of a schema increases as its type frequency decreases. This is graphically represented in Figure 2.2.

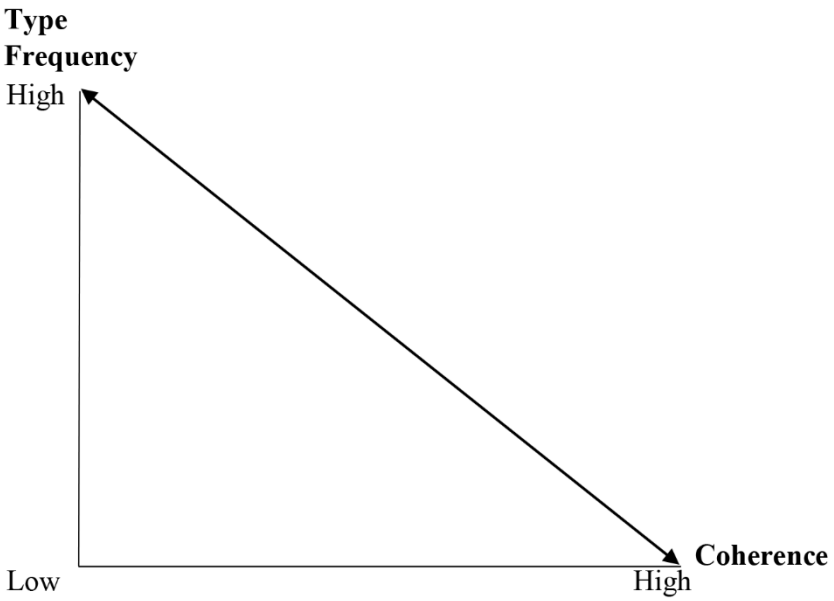


Figure 2.2. The inverse correlation between type frequency and semantic coherence (after Barðdal 2008: 35)

All constructions that are situated relatively close to the cline are considered to be productive; depending on the position on the cline they exhibit varying degrees of productivity. At the leftmost top of the cline, we find open schemas that can be instantiated by a large array of types and that show little internal coherence. Moving towards the right bottom end of the cline, we find constructions that are only productive within a very delimited semantic domain, i.e. constructions with low(er) type frequency and a high(er) degree of internal coherence. Barðdal gives the example of the pattern [*drive* OBJ XP_{crazy}] as in *drive someone crazy*, which is remarkably productive within a specific semantic domain. Speakers can fill the XP-slot with pretty much anything within the semantic domain of 'crazy', e.g. *bonkers, bananas, up the wall(s), mad, insane, over the edge, nuts, wild, batty, to pieces...* At the rightmost bottom end of the continuum, Barðdal argues that token frequency may come in as an important factor *for* productivity, in contrast to what has generally been assumed (cf. supra, Bybee 1985, 1995, Bybee & Thompson 1997, Clausner & Croft 1997). In morphological accounts, item-based analogy is generally perceived of as qualitatively different from rule-based productivity. In a cognitive-constructionist theory that accepts constructions existing at different levels of schematicity, there is no reason to assume an ontological distinction between analogy and productivity – the difference is simply a matter of degree (Barðdal 2008: 91). If one of the members of a schema is highly token frequent, it may come to serve as a model for analogical extensions and attract new members to the schema. This is illustrated in the

work of Zeschel (2012), who focuses on a set of specific collocation clusters in German and English (e.g. *glaring/glühende* N, *glare with/glüh vor* N, *glaringly/glühend* ADJ etc. cf. infra) from a synchronic point of view. Based on the existence of creative extensions of frequent, routinised formulae, he hypothesises that frequently used fixed expressions may come to allow for local extensions and may even give rise to a partially productive constructional schema, provided that speakers form a generalisation over these collocations. As there are certain parallels between Zeschels observations and what we find in the intensifying fake reflexive resultative construction, we will return to his work in somewhat more detail in later chapters. Barðdal (2008: 94-95) suggests that type frequency and token frequency are therefore also in an inverse correlation: much like semantic coherence is an important factor at the bottom end of the cline, "high token frequency should be more important for lower-level verb-specific analogical extensions than for the productivity of high type frequency constructions." Finally, we should address what it means for a construction to not be situated on or relatively close to the cline, but further away from the cline, either in the left bottom area under the cline or in the right upper area above the cline. In the former case, both the type frequency and the internal consistency are fairly low, in which case we can hardly speak of a productive construction. In the latter, a category would have to have both a very high type frequency and a high degree of coherence, which is rather unrealistic, in that we would not expect so many types to belong to the same category. In sum, "the categories relatively close to the cline will show most signs of productivity, while categories further away from the cline will either be low/non-productive (lower leftmost sphere) or non-existing (upper rightmost sphere)" (Barðdal 2008: 39).

The interaction between type frequency and semantic coherence has also been picked up by Suttle & Goldberg (2011), who argue that there are three factors that positively influence productivity, i.e. type frequency, variability and similarity. It is hypothesised that language users are more confident in extending a pattern that displays high type frequency and high variability, in which variability is defined by the *range* of attested types. As the degree of variability positively correlates with type frequency, the two are often confounded. Although both type frequency and variability display independent effects on syntactic productivity, they are also involved in an interaction in that the effect of variability is found to be stronger in cases of high type frequency (i.e. low coherence) – a finding which seems to add strength to the hypothesised inverse correlation between type frequency and coherence in Barðdal's model. In addition, similarity has a positive effect on syntactic productivity: if a new coinage is found to be similar to an already attested item, it is found to be more acceptable (cf. the role of analogy at the bottom of the continuum in Barðdal's model). Similarity also enters into an interaction with variability in two important ways. First, if the similarity between the new and already attested items is very high, language users are *less* confident about the new item if the distribution is highly variable. If there is only moderate similarity, however, high

variability has a positive effect because it shows that the construction is already attested with multiple semantically similar classes. The experimental data lead Suttle & Goldberg to suggest that the complex interactions between the effects can be combined in the notion of "coverage". Rather than looking at the type frequency, variability and/or similarity of the attested instances in isolation, coverage considers the relationship between the new type and the attested types and is thus defined as "the degree to which attested instances cover the category determined jointly by attested instances together with a target coinage" (Suttle & Goldberg 2011: 1254). In a way, the notion of coverage is similar to Clausner & Croft's (1997) definition of the degree of productivity as "the proportion of [its] potential range which is actually manifested". What is important is that the studies of Barðdal and Suttle & Goldberg, although their focus and findings are somewhat different, demonstrate how both (type) frequency and semantic aspects need to be taken into account when studying the productivity of constructions.

2.1.2.3 Productivity in diachrony

From a diachronic point of view, the degree of productivity of a construction, which was defined as the extent to which a construction is capable of attracting new members, is subject to change over time (Barðdal 2008, Hilpert 2013). In the traditional morphological literature, however, the diachronic implications of productivity have largely been ignored: the main application of the frequency-based measures is to compare the productivity of functionally equivalent, rivalling word-formation rules in order to determine which rule is more likely to be applied to new words. A constructionist theory like Diachronic Construction Grammar provides a better framework for exploring shifts in productivity as one of the many constructional changes a construction may undergo. Following Barðdal (2008) and Boas (2008), Traugott & Trousdale (2013: 17) define productivity of a schema as (i) the extent to which it sanctions other, less schematic subschemas and (ii) the extent to which its applicability is constrained by restrictions. Expanding productivity is discussed as one of the three changes, alongside an increase in schematicity (cf. §2.1.3) and a decrease in compositionality (i.e. the transparency in the link between meaning and form), that are often found to accompany the process of constructionalisation (Traugott & Trousdale 2013: 113-116). The authors argue that constructions may undergo an expansion in terms of frequency of use of constructs (i.e. token frequency) and construction-types (i.e. type frequency), but that both issues should be kept apart. Whereas grammaticalisation research has generally privileged increases in token frequency (as a result of host-class expansion, i.e. increase in type frequency), Construction Grammar puts the main emphasis on changes in type frequency (Traugott & Trousdale 2013: 114). Constructions may over time relax the constraints on certain slots and extend their collocational range: if a schema comes to be used with more types, this is regarded as evidence of the extensibility, i.e. productivity of that schema. The approach

of Traugott and Trousdale is mainly qualitative; in their concluding chapter, the "measurability" of historical productivity is explicitly signalled as an area for future research (2013: 238).

Seeing as there already exists an elaborate toolkit for measuring different dimensions of productivity in synchronic data, there is nothing to keep researchers from applying these methods to diachronic data as well. For example, Hilpert (2013) evaluates whether the measure of potential productivity can be used to gain more insight into the productivity evolution of the V-*ment* construction (e.g. *adjustment, refreshment*...) in the history of English. Rather than comparing the potential productivity of multiple constructions at one given point in time, he calculates the potential productivity for one and the same construction at multiple points in time. The results are in line with intuitive expectations: the *-ment* suffix has gradually lost its footing to the point where it is no longer attached to new verbs. Barðdal & Gildea (2015) propose that the hypothesised inverse correlation between type frequency and semantic coherence also makes several predictions about historical productivity. If two or more constructions are competing to attract new items, it is expected that the construction with a higher type frequency will attract more new items than constructions with a lower type frequency. Moreover, if the low type frequency construction is also low in coherence, it may *lose* some of its types to the higher type frequency construction, and eventually fall into disuse. This prediction is borne out by the study of the history of a number of case and argument structure constructions in Germanic languages (Barðdal 2009). Swedish and English have lost morphological case and only have the nominative subject construction in the present-day language: as predicted, the most infrequent construction (i.e. the genitive object construction) disappeared first, whereas the constructions with a higher type frequency (like the oblique subject construction) were able to resist a little longer. In Modern High German, subjects are generally in the nominative case, objects in the accusative, indirect objects in the dative and nominal attributes in the genitive case. The low type frequency constructions with non-nominative subjects have fallen into disuse, leaving only a small number of Dat-Nom or Acc-Nom predicates behind. Icelandic is much more conservative in that only the construction lowest in type frequency, i.e. the Dat-Gen construction, has disappeared and merged with the more common Dat-Nom construction; other low type frequency constructions are maintained for now, but they have been losing verbs to higher type frequency constructions. At the same time, low type frequency constructions are expected to be maintained if they are characterised by a high degree of semantic coherence and/or a high degree of (phonological or formal) similarity among the tokens, which may in turn trigger new analogical extensions. This explains why the dative subject construction in Icelandic has ensured its productivity in spite of a rather drastic decline in type frequency: by reducing its semantic scope, the dative subject construction has increased its semantic coherence in present-day Icelandic, thus safeguarding its extensibility within that particular semantic domain (Barðdal 2008).

In this thesis, we will argue that a truly multidimensional model of productivity, especially when applied to constructions, should include both the frequency-based measures that were introduced by Baayen and colleagues and the semantically inspired concepts (primarily) suggested by Barðdal. A combination of these complementary approaches can lead to new insights on the notion of productivity in all its complexity, both from a synchronic and a diachronic perspective. That is not to say that there are no other factors which may exert an influence on the productivity of a pattern. Baayen (2009: 4) recognises that his productivity measures are based on the idea of "grammar [as] the knowledge of the ideal speaker in a homogeneous speech community", thus acknowledging the fact that there may be idiosyncratic differences between speakers. He also suggests that productivity is subject to stylistic variation, as some morphological categories may be more pertinent in some registers than in others. It has been observed that many derivational suffixes have a higher degree of productivity, both in terms of realised productivity and potential productivity, in written language than in spoken language. Moreover, the productivity rankings of certain affixes may vary across registers: for example, the suffix *–ness* emerged as more productive than *–able* in written language, but in spontaneous conversation, the reverse is the case (Plag et al. 1999). While the impact of these societal or stylistic factors are definitely relevant for the study on productivity, they have generally been dismissed for being too unpredictable (Bauer 2001). As a first step towards including language-external factors in the study on productivity in a more systematic fashion, we will look into regional variation in productivity by comparing present-day Belgian and Netherlandic Dutch (see Chapters 3 and 4). In the following chapters, we will put the multidimensional model to the test by tracking the historical productivity development of one construction over an extended period of time. So far, we have only been concerned with the "overall" productivity of a construction. Nevertheless, a construction may subsume a number of subschemas at different levels of schematicity, which may differ with respect to their degree of productivity. Moreover, changes in productivity may take place at each of these levels in the constructional network. In order to grasp how such local changes can cause rippling effects that affect the entire network, the next section will elucidate how the constructional network can be conceptualised as a taxonomically organised system that is constantly in flux.

2.1.3 A dynamic constructional network

In Construction Grammar, there is general consensus on the idea that all constructions, regardless of their degrees of complexity and schematicity, are stored in a taxonomic network, the structure of which is subject to a number of organisational principles. One way in which semantically and/or syntactically similar constructions can be linked is

through inheritance relationships: constructions that are situated at a lower point in the network inherit certain features from constructions higher up in the hierarchy, while adding their own specific properties as well. It is worth noting in that regard that constructions can inherit features from multiple "parent" constructions (see, e.g., Goldberg 1995, chapter 4 in Hudson 2007 on multiple inheritance of gerunds, or Trousdale 2013 on the role of multiple inheritance in language change). Depending on the specific relation between parent and daughter constructions, Goldberg (1995: 75-81) distinguishes four major subtypes of inheritance links, viz. polysemy links, subpart links, instance links and metaphorical extension links. However, this vertical, hierarchic structure only partly accounts for the interrelationship between constructions in a network. Van de Velde (2014) points out that it is also possible to relate constructions on a horizontal rather than a vertical axis, in which case the form-function relation of a particular construction is (in part) motivated by a neighbouring construction at the same hierarchic level. Horizontally linked constructions thus form a kind of syntactic paradigm, i.e. "a set of alternating forms with related meaning differences" (Van de Velde 2014: 149). Drawing on evidence from psycholinguistic experiments in priming and L1-acquisition, Diessel (2015) as well discusses horizontal links between constructions at the same level of abstraction. Furthermore, speakers can experience a certain familiarity between constructions that exhibit formal or semantic commonalities but which are not strictly speaking in a parent-daughter relationship nor subpatterns of the same parent construction (on such similarity links, see Verhagen 2002, Verhagen 2003a, b on the Dutch *way*-construction and other similar constructions, and Taylor 2004). In a recent article, Pijpops & Van de Velde (2016) have proposed the term "constructional contamination" to describe the phenomenon by which constructions affect other constructions based on superficial similarity links. Such synchronic contamination effects may eventually lead to the merger of constructions which were originally structurally independent (see, e.g., Hilpert 2014, Norde & Strik 2017, and the papers in De Smet et al. 2015 on multiple source constructions). As our main concern here will be with the taxonomic structure of a construction-specific network, i.e. the hierarchy from abstract schema to lexically-specific instance for one particular construction (cf. infra), we will not go into further details here but we will return to these horizontal links in §4.4.

Given the immense complexity of any language, studies on network organisation have shied away from trying to account for the entire network of constructions within a given language and have instead zoomed in on smaller-scale networks involving one specific construction or a set of closely related constructions. To that end, Traugott (2007, 2008a, b) has proposed a descriptive model in which she identifies three constructional levels (also see Trousdale 2012). As is to be expected in a usage-based model, we start out with actual utterances at the bottom; these concrete lexical expressions are generally referred to as "constructs". If a certain number of these lexical expressions are conceived of as formally and semantically similar, speakers may form a generalisation over them in the

form of a "micro-construction". This micro-construction is somewhat more abstract than the concrete utterances, but it is still lexically specified to a considerable degree. Moving up, we arrive at the level of "meso-constructions", which capture the commonalities of the lower-level micro-constructions and are in that respect more abstract. At the highest level, finally, we typically find the "macro-constructions". These schematic patterns are the most abstract in that they generalise over all constructions at lower levels. Traugott later slightly revisits this terminology in Traugott & Trousdale (2013: 16), opting for a three-way distinction between micro-constructions, subschemas and schemas – which is also the terminology that will be used throughout this investigation –, the latter two roughly being the equivalents of meso- and macro-constructions respectively. It needs to be pointed out that these terms are meant to function as heuristic devices enabling a researcher to map out the internal hierarchy of a constructional network, i.e. they should not be interpreted as absolute categorisations. In no way does this terminology imply that we can capture the complex structure of the constructional network in just three discrete levels of abstraction (cf. Hilpert 2013). Rather, constructions are represented along a continuum from schematicity at the highest (schema) level to lexicality at the lowest level. The main advantage of such a Lexicality-Schematicity hierarchy is that it allows the researcher to account for both higher-level generalisations – each lower level inherits certain properties from the higher levels – and low-level idiosyncrasies (Barðdal & Gildea 2015: 27). The idea of the Lexicality-Schematicity hierarchy was first introduced by Croft (2003), who applied it to the double object construction. The DOC may be represented by the abstract schema [SUBJ V OBJ_1 OBJ_2], which is associated with the semantics 'cause to receive by means of V'. At a lower level, it is possible to distinguish between different types of transfer by positing subschemas that cluster together certain verb types, without lexically specifying the verb. For instance, we could posit the verb-class-specific constructions [SUBJ $V_{creation}$ OBJ_1 OBJ_2] associated with the semantics 'cause to receive after creation by means of V' or [SUBJ $V_{permission}$ OBJ_1 OBJ_2], associated with the semantics 'enable to receive by means of V'.[3] However, seeing as how not every verb of permission can be used in the ditransitive pattern, it may be more accurate in this specific case to immediately drop down to the verb-specific constructions [SUBJ *permit* OBJ_1 OBJ_2] or [SUBJ *allow* OBJ_1 OBJ_2], in which the verbs of permission are lexically specified.[4] Finally, these micro-constructions are instantiated in concrete sentences like *Sally permitted/allowed Bob a kiss* (but not **Peter enabled her a hug*). A similar model is presented in Barðdal et al. (2011), who elucidate and visualise the hierarchic structure of the West-Scandinavian ditransitive construction. At the maximum level of schematicity, the construction is

[3] We use a slightly different visual representation than Croft (2003: 57) (i.e. [[SBJ PERMIT.VERB OBJ1 OBJ2]/[enabling XPoss]], following Langacker 1987) – but the interpretation remains the same.

[4] Original: [[SBJ permit OBJ1 OBJ2]/enabling XPoss by permitting]] and [[SBJ allow OBJ1 OBJ2]/enabling XPoss by allowing]] (Croft 2003: 58)

represented as [S-V-O_i-O_d], which covers high-level semantic categories, such as actual transfer, creation, mode of communication, etc. Within these larger semantic categories, we find lower-level verb-class-specific and verb-specific subschemas that may be subject to specific formal or semantic constraints. One example of such a formal constraint is the obligatory reflexivity of the indirect object, which does not pertain to the construction as a whole (unlike in the case of the intensifying fake reflexive resultative construction, cf. infra) but may be relevant at lower verb-class-specific or even verb-specific levels. Iwata (2008) distinguishes between verb-specific and verb-class specific subconstructions in the caused-motion construction, pointing out that the different levels in the hierarchy "serve different purposes" (2008: 36). The verb-specific constructions can handle selectional restrictions that pertain to a specific verb, while the verb-class-specific constructions capture certain formal or semantic regularities of an entire verb class. In most cases, it is not an easy feat to determine at which level in the hierarchy a certain subschema needs to be positioned and, as is convincingly shown by Trousdale (2008a), it may well be necessary to distinguish multiple intermediate levels of abstraction, depending on the construction at stake. In his discussion of possessive constructions in English, Trousdale (2008a: 169-170) presents both the possessive construction and the demonstrative construction as subschemas (or meso-constructions) of the more abstract determiner construction. Before moving on to the micro-construction level, he proposes another intermediary level of abstraction that distinguishes between the prenominal possessive construction (the *s*-genitive) and the *of*-possessive construction. Going down the pathway of the *s*-genitive, we then find formally specified micro-constructions like [[Proper N]'*s* N], which can be instantiated in language use by concrete constructs like *Uncle Tom's cabin*.

The diachronic implication of the taxonomic network conceptualisation – or, in other words, the relationship between constructional inheritance and constructional change – is that the internal hierarchic structure of the constructional network is constantly in flux as links are being reconfigured, new nodes are created and existing nodes marginalise or disappear completely (see, e.g., Colleman & De Clerck 2011, Van de Velde 2011, Colleman 2015, Torrent 2015 for some case studies on growth and loss within the constructional network). Such network-internal shifts typically involve changes in schematicity and productivity, viz. changes in terms of the level of abstraction/specificity and the extensibility of a pattern. As was argued earlier, the subschemas in a constructional network may differ in terms of their productivity: a subschema that generalises over a wide variety of types is considered to be more productive than a schema that can only be instantiated by a limited set of construction types (Barðdal 2008). Schematicity and productivity are tightly interconnected: the fact that the more productive schema is subject to fewer constraints entails that it is more abstract or schematic, and, accordingly, situated at a higher level in the hierarchy (Barðdal 2008, Traugott & Trousdale 2013, Perek 2016a). Changes in schematicity can manifest themselves in different ways. First, as language users form generalisations over specific

instances, new subschemas may emerge and, correspondingly, new nodes are formed in the constructional network. When a new construction type is added to the network, the hierarchic organisation gains in complexity and the schematicity at the highest level of abstraction is increased. This type of constructional expansion is attested in the case of the Dutch *krijgen*-passive (Colleman 2015). In the first half of the 20th Century, the *krijgen*-passive was used with four verb-class-specific subschemas: verbs of paying, complex particle verbs with *toe* denoting prototypical events of actual transfer, verbs of delivering and particle verbs of communication. Over the past 50 years, language users extended the use of the construction to new clusters of verbs, viz. verbs denoting future/conditional transfer, verbs encoding "spatial goal" events and some beneficial and maleficial uses. To the extent that speakers indeed recognise and generalise over these verb clusters, new verb-class-specific subschemas have been added to the network of the *krijgen*-passive. Another case is found in the English *way*-construction. In his study on recent changes in the verb slot of the *way*-construction, Perek (2016a) is interested in the semantic types of verbs that joined the distribution and whether the construction shows signs of semantic expansion. Over time, the manner-sense (i.e. the verb specifices the manner in which the motion is performed) starts to allow for a more diverse set of verbs and, accordingly, has become more abstract: whereas it used to have a clear preference for verbs denoting difficulty of motion (e.g. *edge, tramp, trudge*...), verbs that have a more neutral motion sense, like *pace, run, fly*, *swim* etc., have become more frequent over time. At the same time, Perek shows that there may be changes in the distribution that do not directly affect the schematicity of the construction. The recent productivity of the path-creation sense (i.e. the subject creates a path and moves along it), which was already quite schematic and semantically diverse to begin with, is mainly due to local analogical extensions and new members that are subsumed under low-level schemas. In that sense, they do not necessarily contribute to the schematicity of the more abstract schema.

Second, the existing nodes in a constructional network may shift their position over time. We know that a subschema may be extended to new members as the restrictions on the distribution of an existing subschema are eroded (cf. section 2.1.2). This increase in productivity of the subschema is accompanied by an increase in the level of abstraction (or schematicity) of the subschema, causing it to move up to a higher level in the hierarchy. For example, in addition to the emergence of new subschemas in the Dutch-*krijgen* passive (cf. supra), the communication verb subschema, which was one of the four original subschemas, appears to have slightly relaxed its constraint against non-particle verbs and, in that regard, has become somewhat more abstract (Colleman 2015). In his discussion of the *para*-Infinitive family of constructions in Portuguese ([NP_1 V AP/NP_2 *para* (NP_3) V_{INF}], e.g. *Ela deu mil reais pra mim fazer* 'She gave a thousand *reais* for me to do the job'), Torrent (2015) demonstrates how the emergence of new patterns within the network has both caused the network to expand and has triggered several internal reconfigurations. Starting out with a verb-specific pattern in the 13th Century, the *para*-

Infinitive network has since then witnessed verb-specific patterns become more general verb-class-specific patterns and it has welcomed new verb classes to the family. The former development is an example of specific subschemas becoming more schematic and moving up in the hierarchy, the latter illustrates an increase in schematicity at the highest level of abstraction as a result of new subschemas being recruited to the network. Given the intrinsic relations that hold between all members that engage in a network, such changes are unlikely to happen in isolation: the internal organisation of the network has accommodated to the emergence of new subschemas in the form of a reconfiguration of the inheritance links between the subschemas that were already present. The details of these reconfigurations are too complex to be discussed here, but see Torrent's visual representation of the network throughout the centuries (2015: 180, 196-197, 200, 202). The main claim is that once a new subschema starts participating in a constructional network, it builds inheritance relations with the already existing subschemas. Since inheritance links account for the synchronic relations among constructions, they reflect what kind of motivation relationships or generalisations speakers establish between these constructions, regardless of their historical origins. Torrent (2015: 196, 208) finds, for example, that it is possible to posit subpart and metaphorical links between two historically unrelated subschemas within the *para*-Infinitive network.

The case studies that we have discussed thus far are all concerned with expansion in (some part of) the constructional network, but we may also find evidence of *contraction* or *loss* within the constructional network. An erstwhile productive subschema may retreat to a set of specific collocates, thus dropping down to a lower level in the hierarchy. Eventually, subschemas and entire constructions may even disappear completely, leaving behind a (sometimes substantial) number of lexicalised low-level patterns which attest of its former productivity. An example of the latter is found in the evolution of the V-*ment* construction (Anshen & Aronoff 1999, Hilpert 2013). The suffix *–ment* entered the English language attached to French loanwords, but from the 15th Century onwards, native Germanic words ending in *–ment* started to appear. This is a classic case of schema-formation: the similarity between the many borrowed types led speakers to parse the borrowed words into stem + suffix *–ment,* giving rise to a constructional abstraction and making it possible for the suffix to be productively combined with new native words. However, the productivity of the V-*ment* construction was shortlived: although there is still a rather large group of words with the *-ment* suffix, it is no longer available to create new deverbal nouns in present-day English (cf. section 2.1.2 on the relationship between type frequency and productivity). Hilpert (2013) finds that there is a difference with respect to the *kinds* of coinages that were produced in several periods and distinguishes multiple "types" of the V-*ment* construction, depending on certain formal and semantic criteria. These criteria are clustered together in "productivity islands", which he defines as patterns that were characteristic for a given time period (Hilpert 2013: 158). As he points out, the evolution of the V-*ment* construction is more a story of subschemas rising

and falling in productivity than one of general, linear decline. It is obvious by now that the internal developments of the network cannot be captured by one overarching, unidirectional trend, as a change in one area of the network does not necessarily imply that the entire network is following the same path. Returning to the case study on the semantic specialisation of the DOC by Colleman & De Clerck (2011), we may argue that the English DOC in general has become less productive (and less schematic) over the past three to four centuries. The overarching DOC-schema has lost a number of subschemas – or verb-class-specific constructions, in terms of Croft (2003) – and is now considerably more constrained than it was in the 18th Century. However, the obsolescence of some (peripheral) subschemas has increased the semantic coherence of the DOC and has ensured its extensibility to novel verbs, provided these are compatible with the remaining subschemas. The subschema featuring verbs of communication in particular is such an "island of productivity" in that it shows evidence of a high degree of productivity, attracting all kinds of new instrument-of-communication verbs, as evidenced in the Google examples *He snapchatted me a picture of him* or *She tweeted him a question about rhubarb* (see De Clerck et al. 2011 for a more detailed account of the use of new verbs of instrument of communication in the DOC). These examples highlight the importance of taking all levels in the network into consideration and, accordingly, establishing at which level in the hierarchy a certain change is taking place if we want to track diachronic shifts in the network organisation (cf. supra on schemas, subschemas and micro-constructions).

With the exception of the aforementioned case studies, the diachronic aspects of network organisation have not been a major focus of study in Diachronic Construction Grammar. This thesis aims to provide new insights into the mechanisms that drive constructional network reorganisations and the important role that productivity has to play in this. It expands on existing research by bundling the current knowledge about specific mechanisms like analogy versus higher-order productivity, collocational restrictions and different frequency effects, as well as the functional notion of expressivity (which may turn out to be a very important factor, see section 2.3 below) and by showing how these often work in concert in bringing about diachronic shifts within the network. The focus on both higher levels of schematicity and lower-level, specific domains within the network will also allow us to investigate how several mechanisms may – or may not – be at play simultaneously in different areas within the network and how this affects – or does not affect – the other participants in the network. At the highest level of abstraction, we may expect to find a gradual expansion of the constructional productivity in terms of the (types of) lexical items which can fill the different open slots in the construction. This kind of host-class expansions falls under the broader concept of

diffusional changes, i.e. the gradual unidirectional[5] expansion of the distribution or collocational range of a construction, see De Smet (2013). At the bottom of the taxonomic hierarchy, specific micro-constructions are occasionally spawning innovative variations, which may eventually trigger a higher-order generalisation in the form of a partially productive subschema, impacting the internal structure of the network – but they might as well remain no more than occasional creative analogical extensions. Productivity may also be at work at intermediary levels, causing some low-level subschemas that were originally assigned to a semantic niche of the network to be expanded to entirely new verb classes, and to take a more central or higher place in the network. As a (direct or indirect) result, other (competing) subschemas may be losing ground, but the expansion of one schema should not necessarily happen at the expense of another: several functionally equivalent subschemas may continue to co-exist. The existing studies have already shown that it is often – though perhaps not always – impossible to capture the organisational shifts in the network by positing sweeping generalisations. We propose that it may be valuable to consider the level of concrete exemplars in order to explain certain changes. Whereas some very specific instances of a construction are clearly motivated by a multitude of (inheritance or other) links, there are also isolated cases which can only be explained by referring to item-specific mechanisms (much like the "ad hoc mechanisms" in De Smet 2013: 249). Importantly, shifts in the internal organisation of the network may also be driven by different mechanisms in different stages in the development (De Smet 2013). The logic behind this is as follows: the mechanisms that cause changes in the network always operate on the synchronic representation of the network, e.g. local analogical extensions based on low-level regularities. Each actual change within the network may affect the organisation – or "reshuffle the cards" (De Smet 2013: 8) – in such a way that the mechanisms operating on it, although they themselves have not changed, may lead to a different outcome: in the case of analogical extensions, the regularities on which the analogy is based have changed, which makes it possible for new analogical generalisations to be inferred. This is known as the "analogical snowballing effect" (De Smet 2013, based on Ogura & Wang 1996).

One consideration that has not been widely explored yet in the literature involves the possibility of multiple network representations of one and the same construction. Most of the work on constructional networks has been concerned with mapping out the internal structure of *the* network, or part of it, for a particular construction, but this is usually just one possible configuration of *a* network. For example, the ditransitive

[5] Unidirectionality here means that the expansion of the distribution is continuous, "without any serious fallback or fluctuation" (De Smet 2013: 3). However, that is not to say that individual types may never disappear from the collocational range of a certain construction, as long as this does not endanger the overall consistent expansion. We will see how this fits in with the idea of "waves of renewal" in the domain of intensification (cf. §2.3.1.2) in Chapter 5.

construction has usually been represented as a taxonomy of verb-class-specific subschemas and verb-specific micro-constructions (cf. Croft 2003, Colleman & De Clerck 2011 supra), based on the supposition that speakers generalise over verbs. However, language users are likely to recognise multiple regularities in the usage of a construction; there is nothing that prevents them from making generalisations over the fillers of, say, the OBJ-slots as well. This can lead to variations in the hierarchy of the network: in the case of the ditransitive, we would leave the verb slot unspecified and get object-class-specific and object-specific levels instead. The crucial idea here is that language users do not have to pick just one network, but they can access all of these configurations concurrently. It will be argued that this "multirepresentational" or "multiconfigurational" approach to network structure can clarify specific peculiarities in the usage and productivity of a construction that may seem rather puzzling at first sight. In the following chapters, we will return to these issues and illustrate the complexity of network reorganisations by means of a concrete example: the diachronic development of the constructional network of the Dutch intensifying fake reflexive resultative construction. Given that no existing literature has so far brought serious attention to this construction – in stark contrast with the "regular" resultative construction –, the next section will provide a detailed outline of what we already know about the intensifying fake reflexive resultative from existing research. This section will illustrate what makes this construction in particular so well-suited for a longitudinal investigation of the implications of productivity shifts on the internal organisation of the network.

2.2 Resultative constructions

2.2.1 Resultatives and related constructions

The resultative construction is without doubt among the most studied constructions in a number of different linguistic traditions. Additionally, although the main focus of most existing research is on the resultative construction in Standard English, similar patterns in other languages have received some attention as well. Formally, the (English) resultative construction can be represented as [SUBJ V NP $XP_{AP/PP}$], as in *John smashed the vase to pieces* or *Peter painted the door green*. Semantically, the resultative pattern expresses that the subject causes the object to undergo a change of state, as a result of the activity that is denoted by the verb. In some accounts of the resultative construction, it includes examples denoting a change of location (following the metaphor "states are locations", see e.g. Lakoff & Johnson 1980), but others treat these change-of-location examples under a different, albeit related, construction, viz. the caused-motion construction (see Boas

2003 and Goldberg 1995 for examples of the respective viewpoints). The primary aims of the existing body of research are (i) to provide an explanation for the syntactic licensing of the postverbal NP and the resultative phrase, and (ii) to make predictions about the distribution of verbs in the resultative construction (i.e. which verb-RP combinations are (un)acceptable). Glossing over the details, for which we refer to Boas (2003), we briefly discuss how resultative patterns have been treated in various theoretical approaches below, before we turn to a more elaborate discussion of constructionist accounts.

In formal approaches to syntax, resultatives are mostly treated as secondary or complex predication constructions. In several analyses, the postverbal NP and the resultative phrase are said to enter into a subject-predicate relation and form a syntactic constituent that has generally been termed a "Small Clause" (see Chomsky 1981, Hoekstra 1988, Aarts 1989, Hoekstra 1992, Aarts 1995, Bowers 1997, among others). In lexicalist approaches such as Lexical-Functional Grammar and related frameworks, the focus is on the aspectual properties of the resultative construction, which is taken to express a complex event involving a causing event and a caused event. The presence of a postverbal NP and a resultative phrase is licensed by the compatibility of the matrix verb with such a complex event structure (Rappaport Hovav & Levin 1998, 2001, Wechsler & Noh 2001). Boas (2003) notes that both the formal-syntactic and the lexicalist approaches, with their heavy emphasis on either form or (verb) meaning, fail to account for the entire distribution of resultatives. Constructionist approaches manage to avoid this pitfall because their definition of constructions integrates both syntactic and semantic aspects, as well as phonological, pragmatic and discourse properties. In Construction Grammar, the resultative construction is an example of a fully schematic argument structure construction and, like all constructions, it is taken to be meaningful independently of the words which instantiate its open slots (see Goldberg 1995: 180-198, Boas 2003, Goldberg & Jackendoff 2004, Broccias 2004, Iwata 2006, Felfe 2012, Luzondo-Oyon 2014, inter alia, for different constructionist analyses of the formal and semantic properties of the resultative). That is to say, the syntactic structure [SUBJ V NP XP] has its own constructional semantics 'X_{SUBJ} causes Y_{NP} to become Z_{XP} (by V'ing)', independently of the verbs it occurs with. The schematic verb slot can then be instantiated by a concrete verb if this verb is semantically compatible with the overall meaning of the construction; during this process of "fusion", the construction may contribute additional arguments that are not lexically selected by the verb. In an attempt to do justice to the semantic and syntactic variation displayed by the different patterns that are taken to represent the (English) resultative construction, Goldberg and Jackendoff (2004) have established a taxonomy of resultative subconstructions that share certain important properties but differ in other, more specific respects. For instance, they posit a distinction between selected transitive resultatives and unselected transitive resultatives , based on whether the postverbal NP is selected by both the verb and by the construction, or exclusively added by the construction, respectively. In some cases, a transitive verb may be coerced

to drop its canonical object in favour of the postverbal NP that is selected at the constructional level, compare the examples below:

(12) Selected versus unselected transitive resultative with *to drink*

a. ?The man drank his water empty	selected
The man drank his water.	
b. The boy drank his glass empty	unselected
*The boy drank his glass.	

A special subtype of these unselected transitive resultatives, also called fake object resultatives, is the fake reflexive resultative (the latter term is coined by Simpson 1983). As is shown in examples (13) and (14), the reflexive pronoun is obligatory in that it cannot be omitted or replaced by another object. While resultative attributes in transitive resultatives are usually predicated of objects, the reflexive syntax makes it possible for attributes to be predicated of subject referents (Simpson 1983: 145).

(13) The little child really ate herself full on candy.
*The little child really ate full on candy.
*The little child really ate her mother full on candy.

(14) The author had drunk himself to death by the age of 30.
*The author had drunk to death by the age of 30.
*The author had drunk his son to death by the age of 30.

Such fake reflexive patterns have been signalled as constituting a formal subtype of the resultative construction in several languages besides English (see, e.g., Washio 1997 on relevantly similar patterns in Japanese, Huang 2006 on Chinese, Boas 2003 on German, Kiss 2006 on Hungarian). Now compare the examples above with the sentences (15) and (16) below.

(15) Lisa danced herself **to pieces** with her girlfriends last night.

(16) John was laughing himself **silly** over the look on her face.

Although (15) and (16) display the same syntactic pattern [SUBJ V REFL XP] as (13) and (14), there is clearly something different about their semantics. It is improbable, to say the least, that Lisa literally fell to pieces as a result of her dancing all night long or that John turned silly, in the actual sense of the word, while laughing. Rather than denoting the result of the verbal activity, the bolded elements indicate that the verbal activities of dancing and laughing are performed with a certain intensity or repetition. The intensifying potential of resultative patterns in English has not gone entirely unnoticed in the literature, but the detail in which it is described varies to a considerable extent. Goldberg (1995) mentions in passing that some resultative clauses are used hyperbolically rather than literally (e.g. *He tickled her silly*), but she does not give any (intensifying) fake reflexive examples. She claims that a hyperbolic interpretation is possible because the

resultative phrase encodes a clearly delimited endpoint and receives a non-gradable interpretation ("the patient argument has gone over the edge, beyond the point where normal functioning is possible", 1995: 196). Jackendoff (1997: 552) does give some examples of the intensive force of the fake reflexive resultative in English, viz. *Dean laughed/danced himself crazy/silly/to death/to oblivion.* According to Jackendoff, the pattern in English is hardly productive, which, as we aim to show in this thesis, stands in stark contrast to the Dutch variant of this construction. He adds that these are not resultatives but rather belong to "a family of idiomatic intensifiers that share the same syntax as the resultative". This might explain why such intensifying uses of the fake reflexive are not further developed in Goldberg & Jackendoff's (2004) classification of resultative constructions. In his constructional approach to resultatives, Boas (2003) gives the example of *Dan talked himself blue in the face*, which he does not consider to be a true resultative but an example of a construction with the meaning of 'to overdo an activity'. Although he remains rather vague on the exact status of such intensifying examples, he puts it forward as one of the directions of future research: "These constructions clearly have a distinct idiomatic meaning and must be described and subsequently accounted for on the basis of corpus data" (2003: 319). A more detailed account is found in Peña-Cervel (2016), whose article gives a fine-grained analysis of the resultative construction with the PP *to death*. She discusses the non-literal meaning of the resultative phrase *to death* in both reflexive and non-reflexive resultatives like *She laughed herself to death* or *I loved him to death* as an example of an implicational construction: a hyperbolic reading is triggered by the unlikelihood of the literal-resultative scenario and is "weaved" into the meaning of the resultative argument structure construction. World knowledge, together with textual context play a crucial role in deciding which reading is warranted – especially in cases where both the literal and intensifying reading are, in theory, equally conceivable. Nevertheless, there are some verb classes for which the hyperbolic reading is the default (perhaps even the only) one, e.g. the verbs of psychological state (*amuse, frighten, bore, worry, embarrass*...). One important aspect that is highlighted in Peña-Cervel's study is the emotional side of hyperbole: by using a hyperbolic expression, the speaker has an emotional reaction to which he wants to draw the hearer's attention. The PP *to death* is often used to boost negative emotions like boredom, fear or annoyance but it has been extended to positively connoted situations as well (e.g. *love someone to death*). The relationship between intensity and expressivity/emotionality will be taken up in the next section. The non-literal use of the PP *to death* is also discussed by Margerie (2011), who trails its evolution from resultative phrase to degree modifier in different syntactic configurations, including [NP_1 V NP_2 *to death*] and [NP BE ADJ *to death*]. She discusses a number of transitional constructions in which death goes from being an immediate, direct result of the verbal activity to being construed as a more indirect, potential or future result of the verbal activity. The literal sense of *to death* thus came to be interpreted hyperbolically in some contexts, eventually giving rise to the degree modifier *to death*.

Like Peña-Cervel, Margerie emphasises the expression of the speaker's attitude in the use of *to death* as a degree modifier. She proposes that the subjective nature of hyperbole may well have played an important role in the diachronic evolution of the degree modifier construction. In present-day English, most of the transitional categories have disappeared and there is a rather clear allocation of tasks between the resultative and the degree modifier constructions (see Ch3, §3.3.5 for a more detailed description of Margerie's semantic classification of the different uses of *to death*). Although the degree modifier construction has gradually increased in frequency over time, the resultative interpretation of *to death* remains dominant (for now). In a follow-up study, Margerie (2013) looks further into the relationship between the resultative construction and the degree modifier construction by focusing on the ambiguity of *sick*, in patterns like [NP BE SCARED *sick*] and [NP_1 SCARE NP_2 *sick*]. She finds that, in contrast to the patterns with *to death*, the resultative meaning is much less salient than the intensifying meaning. This does not necessarily imply that the intensifying meaning could not have arisen out of the resultative construction, following the same pathway as *to death*, but Margerie offers a different explanation. Based on the predominance of the degree meaning and the low degree of cognitive salience of the resultative meaning, she suggests that the degree meaning is actually the original meaning of this pattern, calling into question common semantic pathways. The degree modifier construction with *sick* originated by analogical modelling on the already existing degree modifier constructions with e.g. *to death* (which did arise out of the resultative construction), *stiff, silly, rigid*. The resultative interpretation has presumably arisen at a later date, also through analogical reasoning: if the pattern with *to death* exhibits lexical ambiguity with the resultative construction, so can the pattern with *sick*. However, although it is likely that other patterns have followed the path from degree modifier to resultative construction, she treats the case of *sick* as "idiosyncratic" and does not give any additional examples.

Finally, there is also a construction bearing similarities to the fake reflexive, both in terms of its syntactic structure and in terms of its expressive semantics, that has attracted moderate linguistic attention, viz. the so-called Body-Part-*Off* construction [BPOC] exemplified in sentences like *They cried their eyes out, She sang her heart out* or *He laughed his head off* (see Jackendoff 1997: 551, Sawada 2000, Glasbey 2003, Goldberg & Jackendoff 2004: 560, Espinal & Mateu 2010, Kudo 2011, Cappelle 2014, inter alia). There is some disagreement on whether the BPOC is to be conceived as a construction in its own right, or pragmatically inferred, i.e. derived from the literal resultative or caused-motion constructions in specific situations. Kudo (2011), and to a certain extent Sawada (2000), propose that the intensifying interpretation is only arrived at because the literal caused-motion scenario is not a feasible event in the real world (cf. the hyperbolic treatment of *to death* by Peña-Cervel 2016). In contrast, Cappelle (2014: 252, 261) argues that although the literal reading "keeps on lingering in the background", the action-intensifying meaning of the BPOC is not (or, at least, no longer) the product of a general interpretive

mechanism, which implies that the BPOC is a separate form-meaning pairing in the constructicon (see also Jackendoff 1997). An argument in favour of the latter is the fact that it has different aspectual properties from regular (fake object) resultatives (Jackendoff 1997, Glasbey 2003). Whereas regular resultative or caused-motion constructions usually describe an accomplishment or focus on an endpoint, the BPOC describes an activity. The different event types are reflected in the time adverbials that are compatible with these constructions: regular resultatives are found with *in*-adverbials (one has accomplished something *in* X time), whereas BPOCs readily combine with *for*-adverbials (one has been doing something *for* X time). Sawada (2000) and Cappelle (2014) primarily focus on the semantics of the construction, pointing out that, in spite of the apparent productivity of the construction in present-day English, there are some idiosyncratic constraints on the use of the pattern. For one, speakers cannot just put any verb in the construction, and they cannot combine verbs and body parts as they please because there is still a certain conceptual association between the verbal activity and the body part in the postverbal NP. For example, while *cry one's eyes out* sounds very conventional, *work one's eyes out* is much less acceptable. This odd effect is presumably caused by the fact that the obvious link that holds between the activity of crying and the eyes is completely absent in the case of *to work*. In contrast, *work their butt/ass/tail off* sounds much better: given that butt/ass/tail are less "specific", the postverbal phrases including these body parts are more flexible than, e.g., *one's eyes out*. This mix of productivity and convention is also prevalent in the Dutch construction that is at the centre of this investigation, as will be shown in the next section.

2.2.2 The Dutch intensifying fake reflexive resultative construction

2.2.2.1 A construction in its own right

Now that we have given some background information on the English resultative construction and the intensifying variants of some of these resultative patterns, we can introduce the construction that is at the centre of this investigation. In Dutch as well, there are several types of constructions or specific patterns that can be used to boost the verbal activity in one way or another. Cappelle (2011a) states that the Dutch language is "bursting with" possibilities to intensify verbs. For example, the comparative pattern [V *als een* N] (e.g. *fietsen als een gek* 'lit. to cycle like a crazy person' or *koken als een bezetene* 'to cook like a possessed person') can express that someone is performing a certain activity (denoted by V) a lot or intensely. Another intensification pattern involves adding a subordinate clause with *dat* 'that', as in [V*'en dat het niet mooi/leuk meer* is] 'lit. to V to the extent that it is no longer pretty/fun'. In Cappelle (2011b), he adds the pattern [*erop los V'en*] as in *erop los flirten/liegen/fantaseren* 'to flirt/lie/fantasise a lot (lit. to flirt/lie/fantasise on it loose)', in which the element of *los* 'loose' indicates that the

activity is performed without bounds. Other similar patterns are [V *van het* V_{inf}] as in *wenen/brullen van het lachen* 'to cry/roar with laughter' or [*sterven van* NP] as in *sterven van schrik/verveling* 'to die of fright/boredom'. While most of these patterns allow for some variation in the verb slot, they are to a considerable extent lexically specified. In this investigation, we focus on the Dutch intensifying fake reflexive resultative construction, which, much like the English examples that were given in (15) and (16), makes use of the (schematic) syntactic structure of a regular resultative construction in order to convey an intensifying meaning. Compare the examples below:

(17) Literal fake reflexive resultative construction
   a. De man dronk zich dood op vroege leeftijd.
   the man drunk himself dead [...]
   'The man had drunk himself to death at an early age.'
   b. Kinderen eten zich al te vaak ziek aan snoepjes.
   children eat themselves all too often sick [...]
   'Kids all too often eat candies until they feel sick.'

(18) Intensifying fake reflexive resultative construction
   a. De man schrok zich dood toen hij de muis zag.
   the man startled himself dead [...]
   'The man was very startled when he saw the mouse.'
   b. De kinderen lachen zich ziek om opa's mopjes.
   children laugh themselves sick [...]
   'The children are laughing hard at grandpa's jokes.'

Given the obvious commonalities in syntax, the difference between instances of the literal and of the intensifying fake reflexive resultative construction is not always clear-cut. Consider the use of *dood* 'dead' in the following example:

(19) Grootvader werkt zich **dood** in dat stoffig atelier.
   grandfather works himself dead [...]
   'Grandfather works himself to death in that dusty studio.'
   'Grandfather works hard in that dusty studio.'

Since *werken* 'to work' – when done in excess or, in this case, in unhealthy environments – is one of the many activities which can potentially lead to someone actually getting themselves killed, sentence (19) is ambiguous: it can mean that grandfather is jeopardising his life by working in such a dusty studio, but there is an alternative reading in which it just means that grandfather is hard at work in his studio. In such potentially ambiguous cases, the speakers can generally rely on textual context and world knowledge to arrive at the correct interpretation. In Chapter 3 (§3.3.5), we will discuss how we operationalised the surface ambiguity for the purpose of the present investigation. Whereas the resultative phrase in the literal fake reflexive resultative construction is pretty much limited to an AP or a PP, the intensifying construction allows for a great deal

of variation and flexibility in its use, as is illustrated by the variety of intensifiers of different formal subtypes in (20) below.

(20) Examples taken from the SoNaR and KB Delpher corpus.

AP
a. Beide taalkundigen ergerden zich **groen en geel** aan het jargon van bedrijven.
both linguists annoyed themselves green and yellow [...]
'Both linguists were very annoyed by the companies' jargon.'

b. Dat terwijl de vakbonden zich een jaar lang **dood** betoogden tegen mij.
that while the syndicates themselves a year long dead demonstrated [...]
'All while the syndicate violently demonstrated against me for over a year.'

NP
c. Ik heb een vriend die zich in de week **het pleuris** werkt.
I have a friend who himself during the week the pleurisy works
'I have a friend who works very hard during the week.'

d. Flink wat kinderen zijn zich bij sommige scènes **een punthoofd** geschrokken.
quite a few kids are themselves during some scenes a pointy-head startled
'Quite a few children were very startled by some scenes.'

PP
e. Op de planken van de AB-club speelde donderdag de Belgische groep Briskey zich **uit de naad**.
[...] played Thursday the Belgian band Briskey itself out of the seam
'On the stage of the AB-club, the Belgian band Briskey played with fervour on Thursday.'

f. Dan amuseren de kleinkinderen zich **te pletter** op het strand.
then enjoy the grandchildren themselves to smithereens on the beach
'Then the grandchildren enjoy themselves very much on the beach.'

NP+PP
g. 9000 bedienden renden zich **de benen vanonder het lijf** om de klanten te behagen.
[...] ran themselves the legs from under the body [...]
'9,000 servants ran around like crazy to please the customers.'

h. De Colombiaan werkte zich **de naad uit de broek** maar vond zelden steun bij zijn ploegmaats.
the columbian worked himself the seam out of the pants
'The Columbian worked very hard but he rarely got any support of his team mates.'

NP+AP
i. Heeft British Airways het bij het rechte eind? Analisten peinzen zich nog steeds **het hoofd suf**.
[...] analysts think themselves still the head drowsy
'Is British Airways right? Analysts are still pondering hard (about that question).'

j. Wekenlang schrijven de Belgische kranten zich **de vingers blauw** over het proces-Dutroux.
[...] wrote the belgian newspapers themselves the fingers blue [...]
'For weeks, the Belgian newspapers wrote article after article on the trial of Dutroux.'

NP+part
k. De Arabische vluchteling bedelt en slaapt op treinstations. Huilt zich **de ogen uit**. (example from the Internet, rare category in SoNaR/Delpher)
[...] cries himself the eyes out
'The Arabic refugee begs and sleeps in train stations. Cries his eyes out.'

From a purely syntactic perspective, only the examples (a-b) and (e-f), with either an AP or PP intensifier, represent what is traditionally viewed as resultative syntax: the

reflexive pronoun is in direct object position and the AP or PP can be analysed as resultative phrases. In examples (c) and (d), with an NP intensifier, one might recognise the formal structure of a different construction, viz. the double object construction, in which case the reflexive pronoun is analysed as the indirect object and the NP as the direct object.[6] A comparable similarity holds between examples (g-h) and the caused-motion construction, in which the verb is followed by a direct object NP and a prepositional complement (Cappelle 2014). The NP+AP examples in (i-j) appear to be a hybrid construction in between the resultative and caused-motion constructions. Strikingly, the examples with NP+PP and NP+AP intensifiers all involve a body part or piece of clothing, which reminds us of the English Body-Part-*Off* construction. However, it is (k) that represents the real Dutch reflexive counterpart of the English Body-Part-*Off* construction, but it is extremely rare in newspaper data. In the examples (g-k), the reflexive pronoun serves as a kind of possessor dative to indicate inalienable possession, which is why Cappelle (2014: 263) reserves the term "possessor reflexive" for those cases.[7] However, all of this raises the question of whether positing different constructions does not detract from the obvious similarities between the examples in (20), namely that the bolded elements unmistakably boost or intensify the verbal activity, regardless of their syntactic category. That is not to say, of course, that the syntactic category of the intensifier should be disregarded completely; on the contrary, it is not unlikely that intensifiers from different categories exhibit slightly different behaviour, both from a synchronic and a diachronic point of view (see Chapters 4 and 5). We will opt for "intensifying fake reflexive resultative construction" as an umbrella term for the overarching construction and consider all types exemplified in (20) as lower-level, formally specified subschemas.

There are several arguments in support of this pattern being a separate form-meaning pairing within the present-day Dutch constructicon. Some of these were already invoked in the context of the BPOC by Cappelle (cf. supra), but we will now clarify how they can be applied to the Dutch intensifying fake reflexive resultative construction as well. Firstly, the fact that the same "expressive" or intensifying meaning has different formal realisations across languages is invoked as an argument in favour of the existence of a separate construction in the language-specific constructicon (Cappelle 2014, based on

[6] The question whether the reflexive pronoun functions as a direct or an indirect object does not seem relevant in the case of the intensifying fake reflexive resultative construction.

[7] Dutch also has a possessive variant of (g) to (k): one could say *zijn longen uit zijn lijf* ('his lungs out of his body'), *zijn hoofd suf* ('his head drowsy') or *zijn ogen uit* ('his eyes out'), without reflexive pronoun. However, for (g-h) and (i-j), the reflexive construction is by far the most frequent one. Just to illustrate, a quick search in the journalistic subcorpora of SoNaR for *POSS longen uit X* yields only 7 examples, whereas *REFL de longen uit X* gives over 50 hits. Examples like (k) are the only ones in which the possessive variant appears to be more frequent, which is likely due to the influence of the English Body-Part-*Off* construction. English, on the other hand, does not have a formal (i.e. reflexive) equivalent of the subtypes in (c-d), (g-h) and (i-j).

Croft 1998). Croft's (1998) proposal relies on the assumption that linguistic units, including grammatical patterns or constructions, that are pragmatically inferred should have formal equivalents across languages. The absence of such translational equivalents is considered as proof that the unit or construction should be stored separately in each language. Although this assumption has met with some criticism, Cappelle argues that Croft's "test" holds out for languages which otherwise have similar encoding possibilities. Concretely, if the intensifying fake reflexive resultative construction was pragmatically derived from the literal fake reflexive resultative, we would expect languages with formally equivalent fake reflexives to arrive at the same intensifying construction. Although there are some near-equivalents in other languages (cf. the BPOC above and the examples with *Tod/death/mort/muerte* in the next section), Dutch really stands out from the crowd in that it allows for such a wide variety of different intensifiers from different syntactic categories. Secondly, like the BPOC, the intensifying fake reflexive resultative construction describes an activity and is found with time phrases expressing duration (see, e.g., examples b and j above), which are incompatible with a resultative reading. The time phrase can even act as a disambiguating factor in verb-intensifier combinations which could theoretically receive both a literal and an intensifying interpretation: compare *Hij dronk zich wekenlang dood* 'He drank a lot for weeks' versus *Hij dronk zich dood in één dag* 'He drank himself dead in one day'. The third argument is also aspectual in nature: by describing an activity event type, the construction does not denote a clear endpoint, unlike regular resultatives; Goldberg (1995) claims that result adjectives therefore cannot be gradable. This argument does not apply to the intensifying fake reflexive resultative construction: there are several examples on the Internet in which the adjective is "graded", e.g. *Het personeel schrikt zich een beetje rot* 'lit. The staff startles itself a little rotten', *Ik erger me een beetje dood aan Emmanuel Rutten* 'lit. I annoy myself a little dead at Emmanuel Rutten', *Ik lach me een beetje stuk* 'lit. I laugh myself a little broken'. Fourth, there are numerous intensifiers that presumably never had a literal counterpart and which can only be explained if we accept the existence of a separate intensifying construction (cf. the patterns with *sick* in Margerie 2013, supra): in what world can we imagine a scenario in which one would literally 'laugh oneself a monkey' or 'startle oneself a pointy head', see the examples in (20). What is more, it appears that whichever lexical item takes up the function of intensifier, irrespective of its original semantics, the native speaker of Dutch will naturally arrive at the intended intensifying interpretation. This is best illustrated by the fact that the lexical element occupying the INT-slot does not even have to be a real word in the Dutch lexicon. In analogy to a variety of (informal names for) real diseases (*de pleuris* 'the pleurisy', *de tyfus* 'the typhoid', *de tering* 'the consumption', etc.), the construction may host a series of fictitious diseases like *het schompes*, *het apezuur* (lit. the monkey-acid) or *het leplazerus*, as well as many other nonsense words (e.g. *de rambam*, *het habbiebabbie*...). From a constructional perspective, we may claim that there exists an abstract schema [SUBJ V REFL INT], associated with the

semantics 'Subj Vs excessively/intensely'. As intensification has been found to be a very flexible operation, the exact nature of intensification may vary to some extent depending on the context (Bolinger 1972, Claridge 2011, Zeschel 2012). In this construction the intensity may apply to a very specific dimension of the verb and the semantics may require a slightly different formulation. For instance, *zich rot fietsen* may be translated as 'to cycle very fast' and *zich blauw betalen* is more adequately paraphrased as 'to pay a lot of money' than as 'to pay intensely'. The verbal dimensions that can be intensified involve duration, frequency/habituality, number of participants and amount of result (Rainer 2015: 1345 cf. also §3.3.6). In every one of the examples above, the individual lexemes themselves do not *necessarily* have inherently intensifying semantics – although they can have some inherent expressive potential, cf. §2.3.2 on taboo terms. Rather, it is the schematic pattern, i.e. the construction, that appears to contribute this intensifying meaning, a phenomenon that has been referred to as constructional coercion (Goldberg 1995, Michaelis 2002, 2004, Lauwers & Willems 2011, Audring & Booij 2016). It has been argued that coercion ties in with productivity, in as far as a construction is "productive to the extent that it can coerce new words to appear in it" (Suttle & Goldberg 2011: 1238), though of course productivity is not limited to cases of coercion (cf. §2.1.2 supra). In this case we do find that, as the semantics are primarily determined by the empty schema, speakers of Dutch can experiment freely with the fillers of the INT-slot, as was already demonstrated by the creative examples in the introduction (repeated here as (21) to (23)).

(21) Paps bed aan het maken, net doorheen gezakt god schrok me **de tieten van me lijf af**. (25/02/2013)
[...] startled myself the tits off my body off
'[Dad was fixing my bed and fell through.] God, it startled the hell out of me.'

(22) Ik verveel me **de neten** en kan wel gaan leren maar daar heb ik helemaal geen zin in. (19/06/2017)
[...] I bore myself the nits [...]
'I am so bored... I could go and study but I don't feel like it at all.'

(23) Ik was denk ik vergeten dat Florian er stond.... ik schrok me **de knetters** van een karton (15/08/2017)
[...] startled myself the sparks [...]
'I must have forgotten Florian was there... I was so startled by the piece of cart board.'

At the same time, there is a great deal of convention involved in the use of this construction, which downplays the impression that "anything goes": certain intensifiers and certain verb-intensifier combinations occur with a much higher frequency than others. Even on Twitter, a social medium that is known to allow for informal language use and linguistic creativity (see the examples above), we see a certain number of fixed combinations recurring. To illustrate this, (24) below lists the intensifiers and their frequencies found in the first 100 unique hits on Twitter (search query performed on 4 October 2017) of *schrok me* 'startled myself' followed by an intensifier:

(24) *dood* 'dead' (34), *rot* 'rotten' (14), *kapot* 'broken' (9), *de tering* 'the consumption' (8), *een hoedje* 'a little hat' (8), *de tyfus* 'the typhoid' (5), *wild* 'wild' (4), *de tandjes* 'the little teeth' (3), *lam* 'paralysed' (2), *de pleuris* 'the pleurisy' (2), *de graftering* 'fictitious disease' (1), *te pletter* 'to smithereens' (1), *de pest* 'the plague' (1), *te pleuris* 'to pleurisy' (1), *de blubber* 'the blubber' (1), *de kanker* 'the cancer' (1), *het apezuur* 'fictitious disease' (1), *wezenloos* 'blank/vacant' (1), *het leplazerus* 'fictitious disease' (1), *te barsten* 'to bursts' (1), *het apelazerus* 'fictitious disease' (1)

The list contains 21 unique intensifiers, 11 of which are one-offs (including 3 fictitious diseases). While this is certainly indicative of a high degree of productivity, the results also provide evidence for the existence of certain conventionalised collocations: together, the combinations *zich dood schrikken* 'to startle oneself dead/to death' and *zich rot schrikken* 'to startle oneself rotten' account for almost 50% of all tokens. That is, these specific verb-intensifier combinations have become so frequent that they are considered to be conventional ways of expressing that one is very startled. It is precisely this intriguing mix of productivity and conventionality that makes the intensifying fake reflexive resultative construction a suitable candidate for an in-depth investigation of historical productivity and network-internal changes. The preliminary results suggest that, in present-day Dutch, the construction presents a complex constructional network that is made up of a combination of several islands of productivity at different levels in the hierarchy on the one hand, and conventionalised, virtually fossilised collocations on the other.

2.2.2.2 The construction in previous literature

Scant attention has been paid to the intensifying fake reflexive resultative construction in Dutch – or the Dutch resultative construction in general, for that matter (see §30.3.2.3 in de *Algemene Nederlandse Spraakkunst* and Broekhuis 2013: 269-273 on "resultative complements" for a descriptive grammar point of view).[8] Strikingly, some studies that are focused on the more general resultative construction do give some example sentences which we would classify as instances of the Dutch intensifying fake reflexive resultative construction. In as far as such examples have been treated as a somewhat discrete category, only the formal aspects of the (fake) reflexive pattern have received some attention; its intensifying meaning has generally not been recognised as such. Hoekstra's (1988) work on resultatives within the Small Clause Theory contains multiple examples

[8] Note that in Dutch, the name "resultatiefconstructie" may also refer to a different construction, i.e. *Hij krijgt zijn appels verkocht* 'He manages to sell his apples' or *Ze kreeg haar taak niet op tijd afgewerkt* 'She didn't manage to finish her assignment in time'. The combination of the auxiliary *krijgen* 'to get' with the past participle of V means something along the lines of 'to manage to V a person/object' (Landsbergen 2009, Clement & Glaser 2014, Colleman 2015).

of the intensifying fake reflexive resultative construction. In the first list of examples meant to illustrate the properties of the resultative construction, we find *Hij werkte zich suf* 'He works himself full' (his translation), together with more prototypical resultative examples like *Hij liep zijn schoenen scheef* 'He walked his shoes worn on one side' and *Hij schaatste het ijs kapot* 'He skated the ice cracked' (1988: 115-116). Further in the text, he gives examples like *dat Jan zich suf praat (over het weer)* 'lit. that John himself dazed talks about the weather', *dat Jan zich (aan dat onderwerp) een ongeluk werkt* 'lit. that John himself on that subject an accident works' and *dat Jan zich (om die mop) rot lachte* 'lit. that John himself about that joke rotten laughed' (examples and translations by Hoekstra 1988: 127). These are part of a discussion on how a Small Clause analysis can account for the PP between brackets, but once more, they are not set apart from "regular" resultatives. At one point, Hoekstra does briefly touch upon the degree meaning of English examples like *I worked myself to death*, but he adds that

> the mere fact that the particular situation of being dead can be brought about by my working is the basis for the inference that I worked very hard; *the meaning is not different from the meaning of the other examples given* (1988: 121, emphasis added)

He does not specifically point to any Dutch equivalents, although some of his examples clearly prove that Dutch resultatives can also express a degree meaning: *hij werkt zich suf*, *hij praat zich suf*, *hij werkt zich een ongeluk* and *hij lacht zich rot* can be paraphrased as 'He talks/works/laughs a lot/intensely'. Also operating within a formal framework, Everaert & Dimitriadis (2013) explicitly analyse examples like *Hij rent zich rot* 'He runs himself to the ground' and *Hij werkt zich een ongeluk* 'He wears himself out by working' (their translations) as cases of secondary predication, in which the postverbal phrase expresses an undesirable resulting state. However, Everaert and Dimitriadis ignore the fact that these examples could also be paraphrased as 'He ran extensively/intensely' or 'He works hard', in which case the postverbal phrase acts as a booster to the verbal activity. This ambiguity *is* signalled in the section on resultative constructions in the *Syntax of Dutch* [SOD] (Broekhuis et al. 2015: 253-256): it argues that the examples *Jan schreeuwt zich schor* 'John shouts himself hoarse' and *Jan werkt zich suf* 'John works himself drowsy' can be taken literally, but they can also bring about an amplifying effect. In many cases, the literal interpretation is not readily available and the amplifying reading is the default interpretation, as is illustrated by the examples *Jan lacht zich rot/slap* 'Jan is laughing himself silly' and *Jan werkt zich te pletter/uit de naad* 'Jan is working terribly hard' (SOD translations). In spite of this, however, the SOD does not explicitly treat these cases as constituting a separate construction. Similarly, sentences with an NP intensifier like *Hij lacht zich een aap/breuk/ongeluk/kriek* 'He laughs himself silly' are analysed as regular double object constructions with an amplifying effect. The SOD seems to run into difficulties when examples like *Hij lacht zich de tranen in de ogen* 'He laughs like mad' or *Hij schreewde de longen uit zijn lijf* 'He shouted extremely loud' are described as "regular

resultative constructions [which] *confusingly* [...] are also most naturally interpreted with an amplifying reading" (SOD: 255, emphasis added). We believe that our analysis in terms of an overarching intensifying fake reflexive resultative construction which subsumes all formal variants, can take away some of this confusion.

We also find casual remarks on the intensifying fake reflexive resultative construction in studies that are not, or at least not primarily, aimed at an analysis of the resultative construction. There are some examples of the construction in Vanden Wyngaerd's (2001) article on the distinction between atelic and telic events. In the context of deriving telicity from the distribution of time adverbials (i.e. telic events naturally occur with *in*-adverbials, atelic events with *for*-adverbials, cf. supra), Vanden Wyngaerd (2001: 84, 87-88) points out that "intensifying resultatives", both in Dutch and English, are problematic because they have the syntax of regular (telic) resultatives, but they occur with durative time adverbials: e.g. *Ik heb me minutenlang/*in drie minuten rot gelachen* 'I laughed my head off for/*in (three) minutes' and *She worked her butt off for/*in an hour* (=BPOC). To solve this "problem", he proposes that the intensifying effect is created through an interpretive mechanism: if a literal resultative interpretation is not possible, the resultative predicate takes on an intensifying meaning and can be interpreted as an unbounded/atelic event (cf. Kudo's (2011) analysis of the English BPOC, supra). Although the interpretive mechanism is meant to provide an explanation for both English and Dutch, Vanden Wyngaerd admits that there is a certain degree of lexical idiosyncrasy that is not easily explained, e.g. English appears to prefer PP intensifiers, whereas Dutch has a predilection for adjectives. We have argued that both the presence of a durative time adverbial and the differences across languages may actually be arguments in favour of considering the intensifying fake reflexive resultative as a construction in its own right. Audring & Booij (2016) use the constructional idiom – which is defined as a pattern with both open and lexically specified slots – REFL DET X *schrikken* to illustrate the effect of constructional coercion, in which the constructional meaning 'to be startled a lot' overrides the lexical meaning of the specific X-filler, e.g. *Ik schrik me een hoedje/een ongeluk/de tering/de rambam/het apezuur* 'lit. I startle myself a little hat/an accident/the consumption/fictitious word/fictitious word'. However, they only make mention of the verb *schrikken* 'to be startled' and limit themselves to "quasi-resultative" NP-fillers (2016: 620). As we have demonstrated earlier in this section, the possibilities are much wider than that: *schrikken* is just one of the many verbs that can be used in this pattern, and most of these verbs can occur with a wide variety of intensifiers from different syntactic categories. In Chapters 4 and 5, we will see that this "constructional idiom" is actually one of many partially specified subschemas within the complex network of the intensifying fake reflexive resultative construction.

The only study to our knowledge that gives a more detailed account of what we are calling the intensifying fake reflexive resultative is Cappelle (2014). Drawing on a corpus of English and Dutch intensifying argument structure constructions, he shows that a

construction can contain several subpatterns which differ in their degree of productivity: while some subpatterns appear to be very productive and allow for combinations to be assembled "on the fly", there are also conventional combinations that appear to be stored in the mental lexicon (i.e. conventionalised collocations, in our terminology). The English construction that is at the centre of his investigation is the Body-Part-*Off* construction [BPOC], which was exemplified earlier in this section. On the basis of data from the Corpus of Contemporary American English, he zooms in on ten body parts to measure the usage patterns and the degree of productivity of the BPOC and its subpatterns. The data show that *to work* is the most frequent verb in the construction and that the most frequent body part is *ass* (as in *one's ass off*). Looking at type frequency and hapax counts, Cappelle finds some subpatterns to be much more productive than others. The subpattern [V *one's guts out*] has the highest degree of productivity with a proportion of 25% hapaxes, whereas the subpattern [V *one's eyes* out] is the least productive one with a hapax-token ratio of about 5%. In the subpattern [V *one's eyes out*], the verbs *to cry* and *to bawl* already account for 119 out of 129 tokens, which suggests that these are conventional combinations. The high frequency of occurrence of some of these combinations may act as a kind of "blocking effect": the existence of a highly frequent pattern may "discourage" language users to extend the individual items that are part of this combination to other elements (cf. section 2.1.2 for the influence of high token frequency instances on productivity).[9] The attraction between particular verbs and body parts may be explained by referring to world knowledge or "encyclopaedic relatedness", e.g. the subpattern [V *one's lungs out*] naturally occurs with verbs of forceful air expulsion like *to scream* or *to cough* (2014: 269-270). In the contrastive section, Cappelle (2014) discusses several "excessive-semantics" patterns in Dutch, see the examples below with the glosses provided by Cappelle (2014: 262-264).

(25) Het vriest de stenen uit de grond.
it freezes the stones out of the ground
'It's freezing very hard.'

(26) Ze zong haar longen uit haar lijf.
she sang her lungs out-of her body
'She was singing her lungs out.'

(27) We betalen ons blauw.
we pay us blue
'We're paying an awful lot of money.'

[9] Blocking effects also exist in the lexicon. Cappelle (2014: 260) gives the example of *water tower* in English versus *château d'eau* in French for one and the same object. Although *castle of water* and *tour d'eau* are in principle possible alternatives in the respective languages, the use of of these terms is blocked by the existence of a highly frequent term.

(28) Ik lach me rot!
I laugh me rotten
'I'm rolling on the floor laughing!'

(29) Ze zong zich de longen uit het lijf.
she sang her-REFL the lungs out-of the body
'She was singing her lungs out.'

(30) We verveelden ons de tering.
we bored us the consumption
'We're bored to death.'

(31) Ik lach me een bult!
I laugh me a hunch
'I'm rolling on the floor laughing!'

The first example is an intransitive caused-motion construction with intensifying semantics, the second example is a close equivalent of the BPOC, but with a full PP rather than a standalone particle. All other examples are instances of the intensifying fake reflexive resultative construction, as it has been defined in this thesis (cf. supra on why we have opted to take all formal variants together under one umbrella term). With the exception of the NP examples in (30) and (31), Cappelle (2014) claims that the excessive patterns are not very frequent in Dutch (though we will show that they are much more common than assumed). For that reason, he only focuses on the intensifying ditransitive construction (his term for the NP variant) in the corpus-based analyses. Using Google as web corpus, Cappelle searched for a selection of 15 NP intensifiers preceded by the reflexive pronoun *me* 'me/myself' to find out whether the Dutch construction displayed similar behaviour as the BPOC in English. The verb *schrikken* 'to be startled' is returned as the most frequent verb, but the verb *zoeken* 'to search' – although only about ¼ as token frequent as *schrikken* – is the most flexible verb, combining with 14 out of 15 input intensifiers. Much as was found for the BPOC, there is considerable variation in the degree of productivity of the different subpatterns of the intensifying ditransitive. In the subpattern [V + *me het leplazerus* 'fictitious disease'], about 25% of the tokens are hapaxes, whereas the combination of [V + *me een kriek* 'a hump'] is found with the verb *lachen* 'to laugh' only. Again, it is argued that the high frequency of some combinations may prevent speakers from extending the intensifiers in these combinations to other verbs. In contrast to the BPOC, Cappelle does not really find any obvious relationship between the verbs and the intensifiers: it appears as if the co-occurrence is not motivated by any conceptual relationship. Nevertheless, our investigation (see Chapters 4 and 5) will show that the lexical semantics of some intensifiers do play a role in restricting their collocational range.

2.2.2.3 Diachronic development of the construction

The exploratory corpus research presented in Cappelle (2014) already points to a number of interesting issues within the theoretical framework of constructional productivity and network organisations (cf. 2.1.2 and 2.1.3 above). Nevertheless, neither Cappelle nor any of the other studies have considered the construction from a diachronic point of view, leaving us with little knowledge about how this intensifying fake reflexive resultative pattern has developed into such a productive – yet partly constrained by convention – construction. We have argued earlier in this section that there are several reasons to assume that the intensifying pattern has developed into a construction in its own right, but it remains unclear at what point exactly it entered the grammar. Following the Invited Inferencing Theory of Semantic Change (Traugott & Dasher 2002: 35-40), it is plausible that the intensifying interpretation was originally arrived at through a context-specific invited inference, which has developed into a conventionalised pragmatic meaning and has finally been reanalysed as a semantic meaning. However, the aim of this thesis is not to describe the constructionalisation of the intensifying fake reflexive resultative construction or to cover its entire history. Rather, we will zoom in on its more recent history by tracking the changes this construction has undergone between the early 19th Century and present-day. First of all, there is an important pragmatic argument for starting our investigation in the early 19th Century. Given that the intensifying fake reflexive resultative construction is not a very common construction, especially in older stages of Dutch (cf. infra), we zoom in on a period for which larger quantities of machine-readable text are available (see also Chapter 3, §3.1.1). The specific time frame is also motivated by the fact that the construction appears to have undergone some drastic changes over the last two centuries. If we look at the Van Dale dictionary entries of the most frequently intensified verbs, like *lachen* 'to laugh', *schrikken* 'to be startled' or *zich ergeren* 'to be annoyed', we immediately find several examples of the intensifying fake reflexive with different intensifiers (e.g. *dood* 'dead', *rot* 'rotten', *een beroerte* 'a stroke', *een aap* 'a monkey', *een hoedje* 'a little hat', *een breuk* 'a fracture', *een bult* 'a hump', *groen en geel* 'green and yellow', *het apelazerus* 'fictitious disease'...). However, if we do the same in the *Woordenboek der Nederlandsche Taal* [WNT], we primarily find attestations with the intensifier *dood* 'dead', the oldest of which, see (32), dates back to the 17th Century. In addition, there are a number of pre-19th-Century examples that suggest *dood* was already unambiguously used to intensify a small range of verbs at that time:

(32) Och, ach ick **lachmen doodt**, ick kan 't niet langher harden. (1617)
oh, ah I laugh-myself dead, I can it no longer bear
'Oh, I'm laughing so hard, I can't bear it any longer.'

(33) Zonder dat hy **zig** met een handwerk behoefde af te slooven, of voor een Smisse **dood** te **sweeten**. (1757)
without that he himself [...] for a smith dead to sweat
'Without having to wear himself out by doing manual labour, or sweat like a pig in a smithy.'

(34) Ik zou **my dood schamen** indien ik zo met Gods woord omsprong. (1782)
I would myself dead embarrass if I so with God's word handled
'I would be very embarrassed if I trifled with God's word.'

(35) Terwijl de misantropische en heraclitische stoicijn **zich** er **dood** over **ergert**. (1798)
while the misanthropic and heracleitic stoic himself it dead about annoys
'While the mysanthropic and heracleitic stoic is highly annoyed over the subject.'

It appears that *lachen* 'to laugh' was already used with a small set of intensifiers other than *dood* 'dead' before the 19th Century, see (36) to (38). Unlike *dood* 'dead', however, these lexical items were apparently not used with any other verbs before the 19th Century.

(36) Maeckt de menschen soo vol bliischap, datse hen seluen **te bersten lacchen**. (1608)
[...] that-they them selves to bursts laugh
'It fills the people with such joy that they are laughing intensely.'

(37) Hy sturf in haer schoot Een suyckerige doot, Daer sy **haer slap** om **loegen**. (1610-1620)
[...] there she herself weak about laughed
'He died in her lap. A sweet death. She had a good laugh about it.'

(38) Kan men zeggen, 't Leezen maakt melancoliek? Wel, ik zal het wederleggen: 'k **Lach my**, leezend, dikwijls **ziek**. (1781)
[...] I laugh myself, reading, often sick
'Is it true that reading makes one melancholic? Well, let me refute that: I often have a good laugh when I am reading.'

We also performed an exploratory search in the Corpus Literair Nieuwnederlands, which includes Dutch literary texts from the 16th to the 20th Century (Geleyn 2016). The search query was based on a number of frequent intensifiers (*dood* 'dead', *suf* 'drowsy', *rot* 'rotten', *te pletter* 'to smithereens', *kapot* 'broken', including spelling variants) and a set of frequently intensified verbs (all forms of *schrikken* 'to be startled', *lachen* 'to laugh', *zich ergeren* 'to be annoyed', *zich schamen* 'to be embarrassed', *werken* 'to work', including spelling variants) in present-day Dutch, in combination with a reflexive pronoun. Parallel to the findings in the WNT, it appears that in the pre-19th Century, (*half*) *dood* '(half) dead' was the only item which could already be used as an intensifier with a variety of different verbs.

(39) Mijn heil, van u te moeten scheyen, Dat kan onmogelijk zijn, ik wil **my dood** gaan **schreyen**. (1686)
[...] I want myself dead go cry
'My god, being separated from you... That is impossible, I want to go cry my eyes out.'

(40) Nu is hy een Poëet geworden. Foei, ik **schaam my half dood**! (1784-1785)
[...] I embarrass myself half dead
'Now he has become a poet. Shame, I'm so embarrassed!'

Just like in the WNT, we also find some early examples of recurring expressions with *lachen* 'to laugh', like *zich slap lachen* 'to laugh oneself weak' and *zich stom lachen* 'to laugh oneself stupid', see (41) and (42).

(41) Zie daar, zie daar, daar is 't. Ik **lach me stom**! (1619)
[...] I laugh myself stupid
'Look over there, over there, there it is. I'm laughing so hard!'
(42) 't Was al Heer Oratyn uw dienaar. 'k **Lach me slap**! (1695)
[...] I laugh myself weak
'It was sir Oratyn, your servant. I'm laughing so hard!'

Although some of the combinations above may have originated as fixed expressions, it is possible that the bond between verb and intensifier loosened over time, thereby allowing the individual elements to enter into new combinations.

Finally, the focus on the more recent history of this construction is also inspired by its particular semantic properties. In fact, the expressive nature of intensification may be an important factor that has contributed to the diachronic expansion and present-day creativity of the construction. The next section will elucidate how intensification and expressivity tie in with one another, and how this bears on the question why expressive, intensifying constructions, such as the construction under investigation here, are interesting for a study on recent diachronic changes in productivity and constructional network organisation.

2.3 Intensification, expressivity and language change

2.3.1 Defining expressivity and intensification

The idea that expressivity is one of the basic functions of language is widely recognised, and the past two decades have seen an increasing interest in the linguistic means that are available to language users for expressing their emotions (as will be shown by the case studies and references that are mentioned throughout this section). In its broadest sense, linguistic expressivity can be defined as the manifestation of "the self" in verbal communication. Given that virtually all utterances are, in some way or other, speaker-dependent, this general sense does not manage to capture the essence of expressivity as a distinct function of language next to the level of mere description or representation

(Hübler 1998: 4-5). Following most scholars working on linguistic expressivity, we opt for a more narrow definition in terms of the explicit expression of the speaker's emotions and attitudes by means of particular linguistic structures, as opposed to linguistic (objective) "descriptivity". It appears that language users experience a universal need for (hyperbolic) expressivity in their means of communication, which is reflected in the extensive repertoire of linguistic means to convey hyperbole or expressivity (Denison & Hogg 2006: 39, Peters 1994: 271). Expressive patterns are particularly vulnerable to habituation, in the sense that they risk losing their pragmatic salience if they become too frequent. Accordingly, speakers feel a continual pressure to change or manipulate their linguistic expressions in order to stay relevant. Haspelmath (1999, 2000) argues that expressivity – for which he prefers to use the term "extravagance" – is the main factor in explaining the irreversibility of grammaticalisation. Speakers want to be "little extravagant poets" and, in order to improve their social success, constantly introduce innovative expressions (1999: 1057). Although they cannot makes changes to the grammar directly, they can come up with new lexical items to supplant grammatical items. Over time, the linguistic community, which wants to share in the social gain of the original trendsetter, may adopt this innovation. In concert with the frequency of use, the predictability of the form also increases, triggering processes of phonological reduction, routinisation and automation, which are common in grammaticalisation (Lehmann 1995 [1982]), and it loses part of its special communicative effect. The explanation for unidirectionality lies in the fact that the reverse development would imply that language users are consciously trying to be *less* expressive, which, according to Haspelmath, goes against the desire to improve social status. Haspelmath's view on grammaticalisation does not remain unchallenged. Geurts (2000a, 2000b) states that the theory of extravagance does not have any explanatory power over the traditional explanation for unidirectionality. Moreover, Geurts points out that Haspelmath makes certain assumptions on the link between extravagance and social status for which he does not offer sufficient evidence. Geurts adds that the early stages of grammaticalisation do not strike him as being particularly extravagant. De Smet (2012) goes as far as to argue that some grammatical changes can happen *because* they are inconspicuous and take place at the unconscious level. The latter point is also made by Traugott & Trousdale (2013: 125), when they posit that Haspelmath's theory runs into problems for changes which have taken place below the level of social awareness and, therefore, cannot possibly be explained by the need to be extravagant. At the same time, however, they argue that "nonconventionality" – which we believe to be rather akin to Haspelmath's notion of extravagance – is an important property of language (based on Langacker's (1987: 69) quote that "a considerable amount of nonconventionality is tolerated (and often expected) as a normal feature of language use") and plays a crucial role in language change and schema-formation. Perhaps the idea of extravagance is not so much a driving force of grammaticalisation, but the desire to be noticed may well be an important

explanatory factor in certain domains in which language users rely on rhetorical effects (De Smet 2017). This section will show that one of these domains is the domain of intensification.

Before we can move on to the close relationship between the concepts of expressivity and intensification, we need to survey some terminological preliminaries, as the notions of intensification and intensifier occur with different senses in the existing literature. One of the most-quoted works in studies dealing with intensification is the grammar by Quirk et al. (1985), whose conception of intensifiers embraces all expressions of degree modification that indicate some point on the so-called intensity scale. Depending on whether the intensifier scales upwards or downwards from an assumed norm, two large subsets of intensifiers are identified: amplifiers, which indicate a relatively high point on the scale, and downtoners, which generally indicate a lower point on the scale (1985: 589-590). The amplifiers are further subdivided into maximisers and boosters, which respectively denote the upper extreme and a high point on the intensity scale. Within the category of the downtoners, four more subsets are discerned: approximators, compromisers, diminishers and minimisers. It has been argued, however, that the amplifiers, and above all the boosters, are the most promising subset for linguistic research because they show the most fluctuation or versatility and the most "colour" (Peters 1994: 271, Ito & Tagliamonte 2003: 258). In the present study, we will therefore only focus on intensifiers that *increase* the degree of the item they modify. An often mentioned prerequisite for intensification is the idea of gradability, which in general has been assumed to be a distinctive property of adjectives. This is most likely the reason why the large majority of studies on intensification are primarily concerned with adjective boosting (Allerton 1987, Klein 1998, Claudi 2006, Doetjes 2008, Van der Wouden & Foolen 2017, inter alia). However, others have shown that gradability is in fact a property found in other categories as well, such as verbs and nouns, and that, consequently, degree expressions have a much wider use than just modifying gradable adjectives (Bolinger 1972, Quirk et al. 1985, Neeleman et al. 2004, Doetjes 2008, Zeschel 2012). Moreover, it has been observed that, even if the modified item is not gradable in itself, the addition of an intensifier may in fact add, or coerce, a gradable interpretation (Paradis 2008, Rainer 2015, Van der Wouden & Foolen 2017). Perhaps as a result of the adjective-centred bias, the literature tends to place a stronger emphasis on degree *adverbs* compared to other linguistic means of intensification. Nonetheless, there have been some studies tackling more diverse intensifying linguistic structures. See, for instance, Zeschel (2012) on verbal and nominal intensification patterns in English and German (e.g. (a) Int + N: *glowing health* and *sirrende Hitze* 'buzzing heat'; (b) Int + ADJ: *blisteringly fast* and *knackig kalt* 'lit. cracky cold'; (c) Int + *with/vor* + V *to seethe with anger* and *kochen vor Begeisterung* 'boil with enthusiasm'), Hoeksema (2012) on different types of elative compounds (e.g. *spinnijdig* 'very nasty, lit. spider nasty', *doodmoe* 'dead tired', *poepchic* 'very chic, lit. shit chic'), Margerie (2014) on the postpositional use of *awful/terrible/horrible* (e.g. *he is angry awful*)

and Van der Wouden & Foolen (2017) on the specific intensifying construction with *possible/möglich/mogelijk* (e.g. *den bestmöglichen Schutz* 'the best possible, optimal protection'). We will show that, within the framework of Construction Grammar, it may be enlightening to look at the possibility of schematic patterns carrying intensifying meaning independently of the individual lexical items that instantiate them.

2.3.2 The expressive nature of intensification and its role in language change

Up to this point we have seen that the function of intensification is to boost an inherent property or descriptive feature of the modified item, but in doing so, intensifying constructions may also highlight an evaluative feature and convey the speaker's attitude (Vandewinkel & Davidse 2008, Gutzmann & Turgay 2012, Van der Wouden & Foolen 2017). Gutzmann & Turgay (2012) distinguish expressive intensifiers from common degree modifiers, positing that the former express a subjective judgment of the speaker which is not part of the truth-conditional content of the sentence, whereas the latter lack this extra dimension. To illustrate this, they give the example of *sau* 'so' in German:

> Beside raising the degree to which the party was cool in [*Du hast gestern eine sau coole Party verpasst*], *sau* expressively displays that the speaker is emotional about the degree to which the party was cool. (2012: 150)

Waksler (2012) makes a similar distinction, reserving the term "over-the-top intensification" for situations in which the intensification construction explicitly marks subjectivity. This is the case when intensifiers like *super* or *uber* surpass certain syntactic, semantic or pragmatic limits, e.g. in the sentence *The crowd at 222 is super random*, the non-gradability of the adjective *random* is overridden, thus marking it for subjectivity (2012: 23). This subjective, evaluative aspect of some intensifying constructions is the reason why intensification and expressivity are so often mentioned together. What is more, the difference between common intensifiers and expressive intensifiers is not always clear-cut, especially from a diachronic perspective: if their frequency of use drastically increases, expressive intensifiers may gradually shed their expressive force and conventionalise into common, i.e. non-expressive, degree modifiers. If Haspelmath (1999, 2000) is right about the crucial role of extravagance as a driving force in language change, this pragmatic wear-and-tear offers some clarification for the constant innovation and lexical renewal observed within the domain of intensifiers: users continue to come up with new expressive intensifiers as they strive to be noticed for their original (or extravagant) language use. Claridge (2011) mentions hyperbole, which she classifies as a subdomain of intensification, as one of the primary means of linguistic creativity. This

makes intensifiers a rewarding object of study for investigations on variation and change, as is most adequately captured in the following quote by Bolinger (1972: 18):

> Degree words afford a picture of fevered invention and competition that would be hard to come by elsewhere, for in their nature they are unsettled. They are the chief means of emphasis for speakers for whom all means of emphasis quickly grow stale and need to be replaced. [...] As each newcomer appears on the scene, it has elbowed the others aside. The old favorites do not vanish but retreat to islands bounded by restrictions (for example, *precious few* but no longer *precious hot*), and the newcomer is never fully successful and extends its territory only so far. Nothing has quite time to adjust itself and settle down to a normal kind of neighborliness before the balance is upset again.

Several decades earlier, Stoffel (1901: 2) already found that "new words [i.e. intensifiers] are in constant requisition because the old ones are felt to be inadequate", and Robertson & Cassidy (1954: 251) even say that "familiarity has bred contempt in the hearer, and one must begin again to find a new 'strong word'". More recently, Blanco-Suárez (2010) talks about intensifiers as "fashion victims", which may disappear as quickly as they have arisen. Related to this is the observation that some intensifiers are characteristic for certain subgroups within the community and, therefore, may signal in-group membership (Peters 1994, Lorenz 2002, Pertejo & Martínez 2014).[10] This may be another reason why new intensifiers are introduced constantly: as intensifiers spread through the community and become more frequent, not only do they lose their extravagance, they may also lose their function of group identification and are due to be replaced by a new form. Still, the renewal is often not complete, in the sense that the introduction of new expressive intensifiers does not necessarily imply that older intensifiers disappear from the language completely; they may be conventionalised and continue to be used as more neutral degree modifiers, as is the case for English *very*, Dutch *erg* 'very' or German *sehr* 'very', or continue to live on in fixed collocations. D'Arcy's (2015) detailed diachronic study of intensification draws a picture of the history of a number of frequent intensifiers in New Zealand English as a combination of both rapid lexical change and gradual grammatical extension. Her data set contains 11 different types of moderators and 59 different types of amplifiers. Despite the small variety of moderators, their diachronic development is marked by competition and renewal: *a bit* shows clear waves of recycling,

[10] Since Lord Chesterfield's comment in 'The World' (1754) on the use of the "fashionable" intensifier *vastly* by fine country women, the use of intensifiers has been associated with women to the extent that women are claimed to use these hyperbolic expressions more often and may have even played a role in the development of certain intensifiers (Stoffel 1901, Jespersen 1922). At the same time, intensifiers are associated with substandard varieties or colloquial usage (Stoffel 1901, Fries 1940). Nevertheless, these associations are not confirmed by Ito and Tagliamonte's (2003) study of the use of English intensifiers in York, suggesting that the influence of social factors demands further investigation.

peaking in the 1900s, the 1920s and again in the 1960s and *fairly* first decreased in the mid-20th Century, before rising to prominence in the 1970s. The domain of amplification is characterised by a higher degree of creativity and versatility, but only five of the 59 amplifiers occur with a frequency of more than 1% in the entire set, viz. *very*, *really*, *quite*, *pretty* and *so*. A first look at the diachronic development reveals that *very* was the dominant intensifier until the early 20th Century, until it was overtaken by *really* around the 1950s. However, if we also take into account the other amplifiers, the picture that emerges is more complex than "simple" lexical replacement. For several centuries, there was a clear gap between *very* and all other forms, which were extremely infrequent. From the late 19th Century onwards, all other forms – not just *really* – started to increase in frequency, giving rise to fierce competition between the intensifiers. The amplifier *really* eventually won out, but not because it simply replaced *very*; it has had to compete vigorously with the other infrequent intensifiers as well. The trajectory that these less frequent intensifiers have followed is remarkably parallel and the data do not contain any real evidence of fashionable intensifiers that were only used for a brief period of time. Overall, her study reveals that the history of intensifiers (including moderators and amplifiers) is not only characterised by "waves of recycling and renewal" (2015: 484), but there are also longer periods of stability or "stasis". The result is that, at any point in time, speakers have an extensive repertoire of both newly introduced and older intensifiers from which they are able to select the one that best fits their needs in a specific situation. Hoeksema (2005, 2012: 97) rightfully observes that "anything to do with degrees belongs to a part of the grammar where lexical parsimony is valued the least".[11] That is not to say that language users have free reign over their choice of intensifiers, as it has been repeatedly demonstrated that a large number of (conventionalised) intensification patterns display strong collocational relations (Greenbaum 1970, Bolinger 1972, Partington 1993, Vandewinkel & Davidse 2008, Van der Wouden & Foolen 2017). Taking into consideration the drive to be original and innovative, however, these collocational restrictions are occasionally flouted, as the "unexpectedness" and novelty of the unconventional combination add to its expressivity (see Chapter 4 for some discussion and concrete examples of such deliberate overrides in the intensifying fake reflexive resultative construction).

The constant addition of fresh members to the category of intensifiers raises the question of where all these novel intensifiers come from. To be sure, the number of new intensifiers that are actual neologisms is negligible; most hyperbolic expressions have developed out of lexical items which used to have a different function or belong to a

[11] However, the domain of intensification is not necessarily all that different from other categories in that regard. Traugott (2008b: 240) says that "in fact every functional category is likely to be renewed many times, and therefore to have many alternative forms. [..] It appears that having several forms that are very close in meaning for one functional category is useful in negotiating meaning."

different syntactic category entirely. In fact, many items that now function as intensifiers are still used in their pre-intensifier meaning as well. This process has primarily been described in terms of delexicalisation and grammaticalisation (although see King 2013 for an interesting case of intensifier borrowing in Acadian French). Indeed, one important line of research in the domain of intensification focuses on the investigation of the syntactic and/or semantic source categories from which intensifiers have been recruited and the factors that have aided in this transition (see Partington 1993, Lorenz 2002, Vandewinkel & Davidse 2008, inter alia, and Norde & Van Goethem 2014 for an interesting case of degrammaticalisation). It has also been observed that the lexical items which have grammaticalised into intensifiers often had a negative connotation. A possible explanation for this is that negative words receive more attention because they are less entrenched, and therefore more unexpected, than positive words (Jing-Schmidt 2007). In the early stages of development, the original lexical meaning can still persist and limit the distribution of the intensifier to strictly negative contexts; the further the intensifier is on the grammaticalisation (or delexicalisation) cline, the less it is bound by such lexical restrictions and the more frequent it becomes. Lorenz (2002: 144-145) gives the example of the English degree modifier *terribly*, which should be positioned about halfway along the cline: most of the top twenty collocates are still primarily charged with a negative connotation, but there are also three clearly positively connoted adjectives (*brave*, *impressed* and *proud*), suggesting that the intensifier no longer means 'to the extent that I consider terrible'. In addition to the propensity towards negatively connoted source items, there appear to be cross-linguistic tendencies as regards the specific conceptual source domains that are more likely to deliver intensifier candidates. Of particular interest are concepts surrounded by a certain taboo: taboo terms are likely to be noticed by the hearer, but their aptitude for being used as intensifiers also lies in the fact that they show variation among social groups and settings and are particularly subject to change. Widespread linguistic taboos are terms for forces of nature related to folk beliefs (thunder and lightning), religious terms (God, the soul, heaven, hell, salvation, damnation...), sexual terms (fuck, cunt, wanker...) and terms for various bodily excretions (blood, piss, vomit, shit...). One domain that appears to be very popular in Dutch are diseases, mostly diseases that are effectively eradicated in the western world, e.g. *de tering* 'the consumption', *de pokken* 'the smallpox', *de klere* 'the cholera', etc. (Hoeksema & Napoli 2008, Napoli & Hoeksema 2009, Hoeksema 2012, Van der Wouden & Foolen 2017). Without question, the source domain that surpasses all others in terms of being the most popular and well-exploited for expressive purposes is the domain of death (which may explain the special status of *dood* 'dead' as one of the first intensifiers in the intensifying reflexive resultative construction, cf. supra). In English, *to death* can be used to boost both verbs and (predicatively or attributively used) adjectives, as in *hate someone to death*, *be sick to death of X* and *a boring to death lecture*. The adjective *dead* has developed into a degree adverb in combination with certain adjectives, such as *dead simple*. In Dutch, there is an

important group of elative compounds with *dood* as their first element, e.g. *doodmoe* 'dead tired', and German has parallel expressions including *zu Tode* and *tod* (e.g. *zu Tode geängstigt* 'scared to death' or *todmüde* 'dead tired'). In Romance languages, we find similar examples in French (verb + *à mort, mortellement*), Italian (*morto di* + noun) and Spanish (verb + *a muerte*), among others (Margerie 2011, Hoeksema 2012). In Chinese as well, the fear of death – which is very profound in Chinese culture – has inspired a number of death-related intensifiers (Jing-Schmidt 2007). Jing-Schmidt (2007) does not invoke the notion of taboo as such, but proposes that a large group of intensifiers find their origin in negatively evaluated emotions, such as fear, disgust or anger.

A second strand of research is not so much concerned with the origin of intensifiers, but with the influence of sociolinguistic factors. Linguists have studied competing intensifiers or recent additions to the repertoire of intensifiers in specific groups of the population (see, e.g., Stenström 1999, Bauer & Bauer 2002, Pertejo & Martínez 2014). In a number of studies both strands of research come together. For example, in the substandard language of young speakers of Dutch, certain quantifiers have developed an intensifying function, viz. *massa's* 'masses' in Flemish varieties of Dutch (De Clerck & Colleman 2013) and *tig* '-ty' in Netherlandic Dutch (Norde 2006). In contrast, we also find mention of intensifiers that are on a declining trajectory, e.g. *hartstikke* 'awfully', which was an up-and-coming intensifier in Netherlandic Dutch in the 1980s (Schröder 1980: 113), but which appears to have already outlived its fashion (Hoeksema & Korterink 2011). For English, Macaulay (2006) has studied the grammaticalisation of the intensifier *pure* in the speech of working-class adolescents in Glasgow and Ito and Tagliamonte's (2003) and Tagliamonte's (2008) articles on intensifiers in Yorkshire English and Canadian English in Toronto also combine both perspectives. Margerie (2014) takes a somewhat different direction in order to account for the emergence of new degree modifier *patterns*, which she demonstrates by means of two constructions in colloquial American English, i.e. [adj + *awful/terrible/horrible*] and [verb + object + *awful/terrible/horrible*]. Taking a constructional point of view, she concludes that such patterns are the result of intertwining pathways of development, driven by the process of analogisation. Norde, De Clerck and Colleman (2014) have also taken a constructionalisation perspective to study non-canonical intensifiers in Dutch, revisiting their earlier work on *massa's* and *tig* and adding recent findings on *duizend* 'thousand' and *een partij* 'a part'.

We can now turn to how the intensifying fake reflexive resultative construction fits in with all of this. In the previous section, it was explained that the construction is intensifying in that the verbal activity is boosted to a higher degree. Although some specific instances of the intensifying fake reflexive resultative construction may have conventionalised to such a degree that they pass unnoticed, we can still add that the construction in general naturally has an expressive (and subjective) meaning component. The Twitter examples below are arguably rather "extravagant" ways of saying that one is

laughing hard or is very annoyed, and they reflect the speaker's attitude towards the situation (also see the examples (21) to (23) in section 2.2.2).

(43) Ik **lach me de ballen uit m'n corduroy** wanneer de eerste Britse Ben Woldring lachend miljonair wordt dankzij deze nieuwe kansen. (Twitter 03/07/2016)
I laugh myself the balls out of my corduroy pants [...]
'I will laugh so hard when the first British Ben Woldring smiles as he becomes a millionaire thanks to these new opportunities.'

(44) Ik **erger me een hersenschudding** aan tv-commentaar vooral bij tennis. (Twitter 09/09/2016).
I annoy myself a concussion [...]
'I'm so annoyed by the (sports) commentary on television, especially during tennis matches.

Given the expressive power of the construction, the expressivity-intensification framework provides a possible explanation for the productivity and versatility of the construction: language users appear to gratefully make use of the construction in order to demonstrate their linguistic resourcefulness and cleverness. Some of the lexical items that fill the INT-slot do have intensifying or expressive uses outside of this construction (e.g. *dood* 'dead' and *ziek* 'sick' are also found in elative compounds like *doodgemakkelijk* 'dead easy' or *ziek grappig* 'sick funny' and some of the diseases are used in expressive exclamations like *Krijg de/het klere/tyfus/tering/pleuris/schompes!* 'get lost!', see also Ch4, §4.1.1.3), but the majority only receive their intensifying potential as a function of being used in this construction.[12] As soon as the intensifier reaches a certain degree of conventionality or familiarity, a novel, more expressive alternative will take the stage. The older intensifier will either gain the status of a conventionalised degree modifier or it may become increasingly obsolete and eventually vanish (see Méndez-Naya 2003 for an example of the latter development involving the English intensifier *swithe*). Given what we know about network organisation, we can establish that the constructional hierarchy is built out of subpatterns with a lexically specified INT-slot and/or V-slot, which display varying degrees of productivity. The degree of productivity is tied to the collocational freedom of the intensifier or verb: some items may be more restricted in terms of their combinatorial flexibility than others, whether bound by their original lexical semantics or not – then again, as was stated earlier, overrides of these collocational restrictions should come as no surprise. Given the fluctuation within the membership of the intensifier category in particular, we expect to find a rather unstable constructional network. New subschemas arise when, e.g., a new expressive intensifier which may have

[12] As a point of terminological clarification, it must therefore be noted that, although we will refer to them as intensifiers for convenience sake, most lexical items that fill the INT-slot in the constructional schema are not strictly intensifiers in their own right.

been restricted to very specific contexts at the moment of its introduction, begins to widen its collocational scope. As the frequency of use of this subschema increases, the effect of novelty and expressivity is bound to gradually wear off. If, as a result, the intensifier develops into a conventionalised degree modifier, the subschema continues to exist at a certain level in the hierarchy; however, if the intensifier falls out of favour, the subschema restricts its collocational scope and drops down to a lower level, perhaps disappearing from the network eventually. In sum, there are several reasons why the intensifying fake reflexive resultative construction is a good candidate for a diachronic study of productivity changes and network reorganisations.

Chapter 3
Corpus and methodology

This chapter is devoted to the compilation of the journalistic corpus that will be used and to the presentation of the methodological steps that were taken during this investigation. In the first part, we will motivate the choice to work with journalistic data and give some background information on the two existing corpora that are at the basis of our journalistic corpus. On the basis of these combined corpora, we have constructed two data sets that will allow us to investigate the use of the intensifying fake reflexive resultative construction in present-day Dutch, as well as investigate the changes this construction has undergone over the past two centuries. The construction and linguistic annotation of the synchronic and diachronic data sets will be discussed extensively in respectively the second and third parts of this chapter.

3.1 Compilation of a journalistic corpus of 19th-21st Century Dutch

3.1.1 Working with journalistic data

This section motivates why we have opted to work with a corpus of journalistic data, which is perhaps a not so obvious choice. In general, linguistic expressions that have a strong expressive and/or subjective force – which is the case for the intensifying fake reflexive resultative construction (cf. Chapter 2) – may be expected to be more common and show more variation or creativity in informal contexts (see Klein 1998, Lorenz 2002, McCarthy & Carter 2004, Claridge 2011, inter alia). Indeed, most of the synchronic studies on intensification and expressivity in language that were mentioned in the final section of Chapter 2 were based on data from spoken conversation. For a diachronic study of the

Dutch intensifying fake reflexive resultative construction in particular, there are both some compelling practical arguments in disfavour of other genres, like informal or literary texts, and some strong style- and content-related arguments in favour of the journalistic genre.

First of all, there are a number of important practical considerations to take into account. Given that this is a diachronic investigation aimed at tracking the changes within a specific construction since the early 19th Century, it is pivotal that we work with a continuous, genre-consistent corpus that covers the entire period under investigation. This criterion immediately rules out a number of mainly informal genres for which we hardly have any data for older stages of Dutch. These include, among others, new media corpora (i.e. web corpora like CoW or the Twitter corpus) and corpora that contain spoken language. Furthermore, this construction in particular requires the use of a sizeable, preferably digitised corpus. It has a very specific communicative function (cf. Chapter 2) that sets it apart from regular, non-intensifying argument structure constructions. This entails that the intensifying construction is not extremely common – or at least not as common as, e.g., the regular transitive, intransitive, or even the ditransitive constructions. If we want to construct a data set that is of sufficient size to perform statistical analyses, we need a very large corpus. This second criterion explains why literary corpora like the Corpus Literair Nieuwnederlands (Geleyn 2016) did not suffice for the present investigation. Although the literary genre does meet the first criterion, i.e. there is no lack of historical data, the existing literary corpora are much too small for this kind of investigation. Taking the example of SoNaR, the STEVIN Dutch Reference Corpus, the journalistic subcorpora contain over 300 million words, whereas the literary genres only amount to 26 million words. This underrepresentation of other genres is even more poignant for older stages of Dutch, for which only newspaper data are digitally available in large quantities. In sum, only the journalistic text genre meets both the criterion of diachronic continuity and the criterion of corpus size.

Second, we argue that journalistic data are in fact better suited for the investigation of expressive language than might be assumed and that journalists may well have good reasons to use creative language. See, for example, the following quote by Reker (1996: 32) with respect to the use of elative compounds in the media:

> Kranteberichten [sic] worden door het gebruik van dikke woorden beter gelezen [...] en sportverslaggevers hebben er een dankbaar hulpmiddel in voor het aantrekkelijker maken van hun verslag.
> 'News articles reach a larger audience by using "fat words" (i.e. elative compounds) [...] and sports journalists make good use of it to add some flavour to their report.'

Several studies have indeed suggested that the general impression of newspapers as a formal, purely fact-based and dry register, which is in turn associated with colourless language use, lacks nuance. Recent years show an increasing interest in what is known as

narrative journalism (also literary journalism or creative non-fiction), a journalistic subgenre that is characterised by a more literary form of storytelling. Inspired by the practices of traditional narratology, narrative journalism finds a way to draw the audience into a story without losing sight of its prime objective of providing factual, objective information (Van Krieken & Sanders 2016b). Although often framed against the background of an increasingly "paperless" news market in a modern world of digitalisation, the core ideas behind narrative journalism are greatly indebted to a journalistic movement from the 1960s and 1970s (Neveu 2014). The New Journalism movement started in the United States of America in the 1960s, heralded by writers like Tom Wolfe, Truman Capote and Gay Talese who discovered that

> it just might be possible to write journalism that would ... read like a novel [...] that it was possible in non-fiction, in journalism, to use any literary device, from the traditional dialogisms of the essay to stream-of-consciousness, and to use many different kinds simultaneously, or within a relatively short space... to excite the reader both intellectually and emotionally. (Wolfe 1972)

Nevertheless, Roggenkamp (2005) finds that the desire to find a balance between reporting facts and engaging the reader goes back to the late 19th Century and that the term "new journalism" was already introduced by Matthew Arnold in 1887.[13] During that period, fiction began to shade off into journalism, and vice versa, as "most editors and reporters believed, as they still do today, that one could be both entertaining and factual. [...] creating within its pages an ongoing dance between the literary (dramatic, sometimes fictionalised, stories) and the journalistic (factual reportage)" (Roggenkamp 2005: xii). In the tradition of Dutch journalism as well, the practice of using storytelling techniques in news articles has gained a lot of professional awareness in the past twenty years (Van Krieken & Sanders 2016b). The use of storytelling in Dutch news articles is not a recent phenomenon either: corpus analyses by Van Krieken & Sanders (2016a, 2016c) demonstrate that narrative structure has been used by Dutch journalists as an important dramatising technique since at least the 19th Century. What is important for our investigation is that this "new" journalism, whenever its actual origin, is said to have presented the journalist with new possibilities in their writing, thus paving the way for expressivity, originality and dynamism in the language of modern newspapers (Markham

[13] Ironically, Tom Wolfe himself admits that this "New" Journalism might not be so new after all: "I have no idea who coined the term New Journalism or when it was coined. I have never even liked the term. Any movement, group, party, program, philosophy or theory that goes under a name with "new" in it is just begging for trouble, of course. But it is the term that eventually caught on. At the time, the mid-1960s, one was aware only that there was some sort of new artistic excitement in journalism. I knew nothing about what history, if any, lay behind it. I was only aware of what certain writers were doing at Esquire, Thomas B. Morgan, Brock Brower, Terry Southern and, above all, Gay Talese." (Wolfe 1972)

2012). In what follows, we shift our attention from the historical changes in the journalistic *practice/genre* to the *linguistic* implications of these changes, viz. the changes in the newspaper register or, in short, "journalese".

A first strand of research is focused on the so-called informalisation and conversationalisation of journalese in recent times. One of the presumed causes behind these tendencies is reminiscent of the paperless market that stimulated the narrativisation of the genre: in a changing news environment, in which information is abundantly available on a number of media platforms, traditional printed journalism must find a way to deal with these new players in the field and stay relevant (Vis 2011). As a social phenomenon, informalisation means, in short, that the hierarchical systems are disappearing, bringing everyone closer together. In journalism, informalisation manifests itself as "an increase of interactivity between the newspaper and the public, as if they are engaging in a conversation" (Vis 2011: 3). This brings us to the notion of conversationalisation, which is understood as the phenomenon in which the language used in newspapers is adopting properties of spontaneous conversation. An important aspect of spontaneous conversation is the fact that the speakers who are engaging in conversation always have to take their addressee(s) into account. This interaction between two or more conversational partners invokes the notion of "subjectivity" and "(inter)subjectification" (Traugott 1989, 2003, 2010): during the conversation – in this case, between the journalist and his readers – the speaker's (i.e. journalist's) attitudes and emotions are brought to the fore and play an important role in the interpretation by the hearer (i.e. newspaper reader) (Vis et al. 2009, 2012, Vis 2011).[14] The conversational nature of language can be measured by observing the presence of certain linguistic elements that are said to express subjectivity, regardless of the text genre. Among these subjectivity indicators are deictic elements like personal pronouns in the first and second person and specific time and place adverbs, modality markers like modal verbs or adverbs, certain causal connectors and specific sentence types like exclamations and direct questions (see Vis 2011 for an extensive overview of all relevant literature on this subject). Several studies distinguish a separate category of modifiers that express and enhance emotional context, including evaluative adverbs, hedges and other explicit stance-marking adverbs, but also degree adverbs and intensifiers (Scheibman 2002). From a diachronic point of view, then, it can be empirically investigated whether written discourse, and journalese in particular, has undergone "subjectification", "informalisation" or "conversationalisation" by tracking the frequency changes of these subjectivity markers in a delineated corpus. Focusing on American and British newspapers in particular,

[14] Subjectivity is defined here as a linguistic concept, i.e. the presence of the speaker in language through the use of linguistic elements that represent the "self" of the speaker and his/her perspective. It should not be confused with the notion of subjectivity in other fields, in which it is often posited as the antonym of objectivity and may have a somewhat negative connotation.

several diachronic corpus-based studies have observed an overall shift towards a more oral and informal style in the second half of the 20th Century – although not all subjectivity markers are found to have increased over time (see, e.g., Biber & Finegan 2001, Westin & Geisler 2002, Steen 2003, Cotter 2003, Biber 2004, Pearce 2005 for a more detailed account of which linguistic markers have contributed to the style shift). For Dutch newspapers the changes in language and style are not as well-documented as in the Anglo-American tradition, but similar findings have been reported. The studies by Vis, Sanders and Spooren (2009, 2012) conclude that Dutch newspapers have also been gravitating towards a more subjective style in recent decades. In a detailed study that compares Dutch newspapers from 1950 to newspapers from 2002, Vis (2011) demonstrates that several of the subjectivity markers that were mentioned earlier have increased in frequency in newspaper language. However, she also notes that this overall rise in subjectivity is not necessarily caused by the fact that the journalist, as a writer, has adopted a different style. Rather, it is mainly attributed to a distinctive increase in reported discourse in news articles. In reported discourse, the journalist gives the floor to other news sources and it is precisely in those direct quotations that we find subjectively flavoured language. She concludes that the increase in subjectivity is an increase in character subjectivity or source subjectivity, rather than speaker or reporter subjectivity (source-reporter subjectivity coined by Vis 2011, based on the character-speaker subjectivity distinction by Pit 2003)

Another line of research is more specifically dedicated to the use of figurative language in journalistic genres. Steen et al. (2010: 43) argue that news discourse is naturally a rich source of figurative language, building on the idea that metaphor helps us deal with the world around us, and that news plays an important role in shaping the public's beliefs and attitudes regarding that world. There are numerous case studies that focus on the use of very specific metaphors in the press coverage of certain events, such as war and combat metaphors in times of financial turmoil (Kitis & Milapides 1997, Charteris-Black & Musolff 2003) or in sports articles (Charteris-Black 2004), construction imagery in political discourse (Musolff 2010), animal metaphors in anti-immigrant coverage (Santa Ana 1999) and a variety of sports metaphors in the domains of, e.g., politics and education (Howe 1988, Semino & Masci 1996, Offstein & Neck 2003). Krennmayr (2011) compares the use of metaphorical language in British newspapers, spontaneous conversation and literary

fiction and academic texts.[15] She argues that news articles are often the result of careful planning, and should theoretically create more room for the journalist to purposefully play around with language and show off his linguistic cleverness than is the case in, e.g., spontaneous conversation (Krennmayr 2011: 143, 150). The results of her study corroborate the common assumption that there are important differences in the way different registers use linguistic phenomena such as metaphors, but they also demonstrate that other factors come into play and that one cannot summarise these register variations in quantitative terms of "more" or "less" metaphorical expressions. A similar study for Dutch, in which the language in Dutch newspapers is compared to the language of conversations, has shown that different registers make use of different *types* of metaphors (Pasma 2011). Additionally, Pasma observes that the language in newspapers shows greater lexical diversity and contains quite a few unique, unconventional deliberate metaphorical expressions.

The subtleties of journalese are too complicated to be discussed in more detail here; naturally, there is some inherent variation within the newspaper register, with some subregisters being more receptive towards journalistic (and linguistic) freedom than others. The key point to take away from the current section is that newspapers overall do contain quite varied, subjective and figurative language. The question now presents itself as to what all of this means for our investigation. First of all, the intensifying fake reflexive construction arguably has a non-literal meaning; at the end of Chapter 2 (§2.3), we also demonstrated in what respect the intensifying fake reflexive resultative construction could be interpreted as an expressive and subjective construction. With what we have learnt in this section, we can now conclude that the journalistic genre is indeed well-suited for the present investigation. The journalistic corpus that will be used in this study is based on two existing corpora, SoNaR and Delpher, which will be briefly described in the next subsections.

3.1.2 SoNaR

The idea for the SoNaR corpus originated by the need for a large reference corpus of contemporary written Dutch. To achieve that aim, the joint Dutch-Flemish STEVIN

[15] Both Krennmayr (2011) and Pasma (2011), cf. infra, adopt a broad definition of metaphor, following the Metaphor Identification Procedure VU University Amsterdam (MIPVU) (Steen et al. 2010). Lexical units are tagged as metaphorical if they are used indirectly, i.e. if the contextual meaning is sufficiently different from the basic meaning, and may be explained by some kind of underlying cross-domain mapping. For metaphorical expressions that are explicitly intended to create a certain rhetorical effect, they use the term "deliberate metaphor". Deliberateness should not be confused with unconventionality: many deliberate metaphors are in fact conventional, but the journalist can also create new, unconventional deliberate metaphors (Krennmayr 2011: 160).

programme, established in 2004, was tasked with building a 500-million-word reference corpus for Dutch that includes texts published in Flanders and the Netherlands from 1954 onwards and is well-balanced in terms of text genres. The corpus comprises no less than 38 text types, ranging from very formal text genres like policy documents and legal texts over newspapers and books to extremely informal genres like SMS, chats and discussion fora. The entire corpus was automatically lemmatised and PoS-tagged with a tagger-lemmatiser software, a large portion of which was subsequently checked and corrected manually. A small subset of about 1 million words was also assigned syntactic and semantic annotation, but we will not go into the details here (Oostdijk et al. 2013).

Since 2014, the full SoNaR corpus can be accessed online at OpenSoNaR, hosted by the CLARIN INL Center. On this online platform, the user can explore corpus distributions, view statistics of specific subcorpora (e.g. frequency lists, vocabulary growth, word clouds...), retrieve n-grams of subcorpora and, most importantly, search the entire corpus. There are different types of search possibilities, depending on whether one wants to search in the entire corpus or in specific subcorpora, or whether one is interested in finding all occurrences of a single word (in its surface or lemmatised form) or a multi-word phrase. We will be using the extended search option that allows us to apply meta-data filters, i.e. text genre and national variety, and to search for regular expressions in CQL (Corpus Query Language). The results of each search query can be exported to external processing or spreadsheet software, although the export function is limited to the first 50,000 hits.

For this investigation, we will only be working with the collections "Newspapers" and "Periodicals & Magazines", which we have taken together as the journalistic subset of the SoNaR corpus, amounting to 305,613,315 million words of running text. The collections consist of both quality and popular national newspapers and periodical newsmagazines: it includes the newspapers *De Standaard, De Morgen*, *Het Laatste Nieuws*, *Het Nieuwsblad* and the magazines *Knack, Knack Weekend, Trends* and *DM Magazine* for Belgium and the newspapers *NRC Handelsblad, Algemeen Dagblad, Nederlands Dagblad, Trouw, De Volkskrant* and the periodical *De Groene Amsterdammer* for the Netherlands. The merged journalistic subset covers the period from 1994 until 2011.

Table 3.1. Word count of the journalistic collections in SoNaR

VARIETY	NEWSPAPERS	PERIODICALS & MAGAZINES	TOTAL "JOURNALISTIC SONAR CORPUS"
Belgian Dutch	152,840,171	79,642,513	232,482,684
Netherlandic Dutch	59,538,177	13,592,454	73,130,631
TOTAL	212,378,348	93,234,967	305,613,315

Table 3.1 shows that the Belgian Dutch part of the corpus is approximately three times the size of the Netherlandic Dutch part, which is something that will need to be taken into account when performing frequency-based analyses in later chapters. This

journalistic corpus of present-day Dutch will serve as the point of comparison for the diachronic journalistic corpus, Delphcorp, which is the subject of the next section

3.1.3 Delpher

Over the past few years, the Koninklijke Bibliotheek van Nederland [Dutch Royal Library, henceforth abbreviated as KB] has developed a number of services and programmes to stimulate digital humanities research in its vast collections.[16] One of these projects is Delpher, a growing digital database that currently consists of over 1.3 million Dutch newspaper issues, some 320,000 books, 1.5 million radio bulletins and 1.5 million pages from periodicals, available via www.delpher.nl. With a main collection of about 11 million individual pages published between 1618 and 1995, the KB has the largest collection of Dutch newspapers worldwide. The collection consists of newspapers published in the Netherlands and the former colonies in the Dutch East Indies, the Americas, the Netherlands Antilles and Suriname.[17] The online platform offers an interactive search module, but the search possibilities are rather limited and do not offer the possibility of searching for complicated regular expressions, for instance. Moreover, the database does not provide information about the total number of words included. Within the scope of this investigation, we have obtained permission from the KB to harvest the Delpher server for all issues of a selected number of newspapers that were available at the time (i.e. July 2016). With the help of Guy De Pauw at Textgain (www.textgain.com), we extracted a demarcated, "countable" subset of full-text data that enables us to search the corpus via the concordance software Wordsmith Tools, version 6 and to use more advanced search queries than are possible in the online search module. There was too little data available for the first decade of the 19th Century, so we started selecting newspapers from the 1810s onwards until the 1990s. The following 12 newspaper titles were selected:

[16] In 2015 and 2017, the KB organised two symposia on historical newspapers as "big data" to give researchers working with its collections the opportunity to share their findings and experience with fellow researchers. The summaries of these symposia are available online at https://www.kb.nl/nieuws/2015/historische-kranten-als-big-data and https://www.kb.nl/nieuws/2017/historisch-onderzoek-in-digitale-kranten-verslag-van-het-big-data-congres

[17] A full overview of all available newspaper volumes can be found on the KB website at https://www.kb.nl/sites/default/files/docs/selectie-1618-1995-alfabetisch.pdf

Table 3.2. The 12 newspapers that were selected from the Delpher collection

Newspaper title	Available volumes[18]	Number of newspaper pages	Total word count in selection
Leeuwarder courant	1813-1942 (129)	26,832	692,211,889
Algemeen Handelsblad	1828-1940 (112)	52,115	1,857,931,523
Middelburgsche courant	1816-1928 (112)	15,607	173,789,378
De Tijd: godsdienstig-staatkundig dagblad	1846-1958 (112)	35,909	1,007,953,455
Nieuwsblad van het Noorden	1888-1994 (106)	28,034	1,479,051,639
Leeuwarder courant: hoofdblad van Friesland	1892-1995 (103)	14,382	951,138,634
De Telegraaf	1893-1994 (101)	52,947	3,432,819,927
Limburgsch dagblad	1918-1994 (76)	21,431	1,001,546,716
De standaard	1872-1944 (72)	9,547	172,770,802
Rotterdamsch nieuwsblad	1878-1944 (66)	16,709	732,595,110
Het vrije volk: democratisch-socialistisch dagblad	1929-1995 (66)	14,997	962,277,270
Het nieuws van den dag: kleine courant	1870-1914 (44)	1,765	654,656,321
Total	1813-1995 (182)	290,275	13,118,742,664

For the sake of continuity, we gave priority to newspapers which covered a large part of the period under investigation. We also included a number of national and regional newspapers that were published during a more limited period of time, but which are well-represented in terms of total newspaper issues and word count. This gives us an enormous corpus of over 13 billion words of running text. Even with the full-text files available, Delpher still has some drawbacks as a linguistic corpus. Unlike the SoNaR corpus, the texts in Delpher are not PoS-tagged or otherwise enriched with linguistic annotation. A corpus of raw, unformatted text complicates the formulation of precise search queries and requires extensive manual post-processing. Second, the digital texts were obtained by running Optical Character Recognition-software [OCR] on scanned newspaper pages. Given the primary focus on mass digitalisation, this process was fully automatised and the results have not (yet) been manually verified. Depending on the printing quality of the original newspapers, the OCR-accuracy shows a lot of variation across different newspapers and decades; especially for the earlier 19th-Century data, the quality is sometimes so poor that entire chunks of texts are illegible. Some common OCR-

[18] This column gives the entire period that is covered by the newspaper volumes, but there are some missing volumes or lacunae in the collection.

mistakes – e.g. individual characters being recognised as aligned pairs, *rn* instead of *m* or *ln* instead of *h,* or vice versa – can be accounted for by using wildcards in the search query. The possible pitfall of using wildcards, however, is that they may generate a lot of extra noise, which is why we will only be using the technique in very specific words (cf. infra). Regardless, the sheer *size* of the corpus makes up for these drawbacks: no other text collections of comparable size are available, making it the ideal corpus for longitudinal research into diachronic changes of relatively infrequent linguistic phenomena of Modern Dutch.

As it would be impractical to work with such an immense amount of data in all stages of the investigation, we have sampled a smaller corpus for the purpose of this investigation which we will be referring to as Delphcorp. The target size was set to approximately 300 million words per decade, sampled randomly from the available newspaper volumes; for the earliest decades that did not contain 300 million words, all data were included, see Table 3.3.

Table 3.3. Contents of the sample corpus Delphcorp

Decennium	Total word count	Selected sample in Delphcorp
1810-1819	7,682,576	7,682,576
1820-1829	18,852,453	18,852,453
1830-1839	51,263,039	51,263,039
1840-1849	75,746,702	75,746,702
1850-1859	115,054,355	115,054,355
1860-1869	146,818,406	146,818,406
1870-1879	306,204,786	306,204,786
1880-1889	493,138,868	299,069,934
1890-1899	878,485,695	300,829,920
1900-1909	1,114,917,563	304,645,355
1910-1919	1,195,721,544	294,708,516
1920-1929	1,240,491,458	304,471,252
1930-1939	1,333,288,611	303,509,962
1940-1949	443,369,131	295,358,332
1950-1959	713,314,677	303,649,139
1960-1969	1,045,329,069	304,995,899
1970-1979	1,626,084,127	299,362,695
1980-1989	1,755,048,499	296,711,421
1990-1999	669,882,042	297,040,183

As we are still dealing with a considerable amount of data, we decided to only focus on every other decennium to begin with – see the decennia marked in grey in Table 3.3. If the exploratory analyses were to reveal significant changes with respect to certain aspects of the construction, viz. frequency, slot fillers and collocations, between two consecutive decennia, the intervening decennium will be included in the more detailed

examination. The next section will explain the procedure that was followed to construct the synchronic and diachronic data sets that will be used for the analyses in Chapters 4 and 5.

3.2 Construction of the data sets

We start out with the intent to construct a comprehensive data set of all instances of the intensifying fake reflexive resultative construction attested in the present-day and diachronic corpora. For our search query, we first turn to the reflexive pronoun, as this is the only element in the [SUBJ V REFL XP] pattern that can be reduced to a finite list of items, see (i).[19]

(i) *me, mij, mezelf, mijzelf, my, myzelf, myself, je, jou, jezelf, jeself, jouzelf, jouself, zich, zichzelf, sich, sichzelf, sichself, sickzelf, sickself, zig, zigzelf, zigself, ons, onszelf, onsself, jullie*

The drawback of only using the reflexive pronouns – most of which are also used as personal pronouns – as input for the corpus search is that such a broad search query inevitably generates a great deal of noise which requires intensive manual post-processing: running the query in the journalistic collections of the SoNaR corpus alone already yields over 2 million results to go through. The best option to keep the number of hits within reasonable limits is to predetermine a number of verbs or intensifiers that can be lexically specified in the search query, in addition to the reflexive pronoun. However, limiting the query to a fixed set of lexical items would drastically reduce the chance of finding creative, unexpected instances of the construction, which will be of crucial importance in our investigation. We therefore opted for a multi-step, cyclic search procedure that allows us to construct an exhaustive (albeit not *maximally* exhaustive) data set of both conventional and unconventional uses of the intensifying fake reflexive resultative construction. The retrieval of all relevant occurrences of the construction in both corpora and the decisions that were made during the process will be outlined below. The differences in the design of the two corpora (cf. §3.1.2 and §3.1.3) required slightly

[19] The spelling variants *my/myself/jeself/jouself/sich/sichzelf/sichself/sickzelf/sickself/zig/zigzelf/zigself/onsself* are based on the 19th and 20th-Century citations in the *Woordenboek der Nederlandsche Taal* [WNT]. These are no longer used in the present-day SoNaR corpus but are likely to get some hits if we go back to older stages of Dutch in the Delpher corpus. In order to keep our search query constant across the entire investigation, they are included here as well. Not included are the forms *hem(zelf)* 'him(self)' and *haar(zelf)* 'her(self)', because their use as a reflexive pronoun in Belgian Dutch (as in, e.g., *Ze wast haar elke ochtend* 'She washes herself every morning') is considered colloquial and is unlikely to occur with any frequency in journalistic data.

different approaches in the search procedure, so the compilation of the synchronic and diachronic data sets will be discussed separately in paragraphs 3.2.1 and 3.2.2, respectively. Both data sets were subsequently annotated for a set of linguistic variables which will be introduced and exemplified in the final section.

3.2.1 Synchronic data set: SoNaR corpus

3.2.1.1 Round 1: [REFL INT] and setting the selection criteria

In the first round of the cyclic search procedure, we queried the journalistic subcorpus of SoNaR [henceforth just SoNaR, for convenience sake] for the reflexive pronouns in (i), followed, within a span of five words, by one out of a previously delineated set of lexical items that may function as an intensifier in the construction. This set is based on occurrences of this construction in the existing literature, i.e. the 15 nominal intensifiers in Cappelle (2014) and the 80 intensifiers listed in Van Beveren (2015), as well as 112 intensifiers mentioned on the blog Pelikanenschurft.[20] We also kept track of all the examples encountered on the Internet and in the media (between October 2014 and October 2015) and added these to the fold. After having removed all overlapping intensifiers, we slightly edited some of the intensifiers and added some wildcards (represented by *) to take into account formal variations or spelling mistakes (the full list of 171 unedited input intensifiers can be found in Appendix III-1).[21] The result was the following list of input items:

(ii) The input items for the search query in round one

adjectives: *blauw, barstensvol, belazerd, beroerd, bewusteloos, blind, dood, doof, failliet, gek, groen, half*, kapot, klem, krom, kreupel, lam, laveloos, lens, levenloos, mal, ongans, plat, purper, raar, rond, rot, scheel, schor, slap, steendood, stijf, stuk, suf, verrot, wezenloos, wild, ziek, zot*

de + noun: *de bagger, de ballen, de benen, de beroerte, de bibbers, de blaren, de blubber, de breuk, de buik, de hell, de joeperdepoep, de klaplaz*rus, de klere, de knetters, de kolere, de knopen, de krampen, de laz*rus, de longen, de mikmak, de naad, de neten, de ogen, de oren, de pelikanenschurft, de pest, de pestp*, de pieperdepiep, de pispleur*s, de pletter, de pleur*s, de pokken, de rambam,*

[20] https://pelikanenschurft.wordpress.com/ (last accessed on October 20, 2017)

[21] For example, some multiword NP or NP+PP intensifiers were shortened to the article and the head noun, e.g. *de benen uit het lijf* 'the legs out of the body' was entered as "de benen" to allow for the retrieval of variants like *de benen vanonder het lijf* or *de benen van het lijf*. As far as wildcards go, an entry like "l*plaz*rus" covers the forms "laplazerus", "laplazarus", "leplazerus" and "leplazarus", as well as any other (minor) spelling variants we might not have thought of. Furthermore, "*t" and "*n" were used to find examples of syncopes like *'t* (*het*) or *'n* (*een*) + noun.

*de schomp*s, de schrikschrak, de shit, de stuipen, de tandjes, de tering, de tierelier, de tieten, de touwtjes, de tranen, de ty*us, de vanalles, de vellen, de vinketering, de wimwam, de ziekte, de ziel, de zolen*

het + noun: **t ap*laz*rus, *t ap*zuur, *t habbieb*bie, *t hellybelly, *t hoedje, *t konijnenzout, *t lamlaz*rus, *t laz*rus, *t lebbes, *t l*plaz*rus, *t ongans, *t ongeluk, *t ongelukje, *t pleur*s, *t rambam, *t schomp*s, *t snot, *t vel, *t wigwam, *t vuur, *t zuur*

een + noun: **n aap, *n apenstaartje, *n beroerte, *n biet, *n breuk, *n bult, *n coma, *n delirium, *n deuk, *n ei, *n gat, *n gil, *n gluut, *n hart*, *n hoed, *n *hoedje, *n hoedjesverzakking, *n hoofddoekje, *n indigestie, *n kriek, *n oelewapper, *n ongeluk, *n petje, *n pizza, *n pleur*s, *n puntmutsje, *n roes, *n rolberoerte, *n rotje, *n sableye, *n sjaaltje, *n stuip, *n tulbandje, *n veertje, *n verlepping, *n voetje, *n *ziekte*

other: *hoedjes, in het zweet, *n slag in, *n stuk in, scheurbuik, slagen in, te barsten, te pletter, tot barstens, tot huilen, tranen*

This search query returned a total of 23,382 hits in the SoNaR corpus. At this point the corpus results still contained a great amount of unwanted hits that had to be manually inspected and filtered out. First of all, all sentences that occurred multiple times in the results – i.e. so-called "doubles" – were omitted from the set. The majority of the sentences that were weeded out in the process did not instantiate the construction we were looking for, like (45) to (47) below, in which the co-occurrence of a reflexive pronoun and one of the input items is merely coincidental.

(45) Wie naar de verwonderde gezichten van de omstanders op de Gras- en Korenlei kijkt, voelt **zich** ook meteen **een aap** in de zoo. (SoNaR-BE)
[...] feels himself also immediately a monkey in the zoo
'Whoever looks at the surprised faces of the spectators on the Gras- and Korenlei, immediately feels like a monkey in the zoo.'

(46) Als Kees je mocht, ging hij voor **je** door **het vuur**. (SoNaR-NL)
[...] went he for you through the fire
'If Kees likes you, he would go through hell for you.'

(47) Begraven worden is nog zoiets dat **me de stuipen** op het lijf jaagt. (SoNaR-BE)
[...] that me the fits on the body drives
'Being buried is another one of those things that makes my flesh creep.'

In addition, there were a number of sentences that, at first blush, might be taken to represent the fake reflexive resultative construction, but which upon closer investigation actually do not, such as (48) to (51) below.

(48) De gangsters geraakten het gebouw binnen maar **beten zich de tanden stuk op** de brandkoffer. (SoNaR-BE)
[...] bit themselves the teeth broken on [...]
'The gangsters made it inside the building but were halted by (or unable to crack) the safe.'

(49) De hypotheekmarkt wordt voortdurend ingewikkelder en toch blijven de meeste kopers of bouwers **zich doodstaren op** de laagste rente. (SoNaR-BE)
[...] the most buyers or builders themselves dead-stare at [...]
'The market keeps complicating, but still most buyers or builders are only focusing on the lowest interest.'

(50) Wat de match tegen de Polen betreft, moeten we **ons** niet **blind staren op** hun zege tegen Portugal. (SoNaR-BE)
[...] must we ourselves not blind stare at [...]
'Regarding the match against Poland, we should not focus too much on their victory against Portugal.'

(51) Al bij hun aankomst **keken** die jongeren **zich de ogen uit** , toen ze hier groot en klein op de fiets zagen rondrijden. In Qatar is dat ondenkbaar. (SoNaR-BE)
[...] looked the adolescents themselves the eyes out [...]
'Upon arrival, the adolescents were amazed by the sight of all sorts of people riding around on their bikes. In Qatar, that is unthinkable.'

These examples are without question formal instantiations of the [SUBJ V REFL XP] pattern. Still, the sentences above were not included in the data set because they do not (or, at least, no longer) qualify as productive, on-the-fly combinations of a verb and an intensifier. It appears that the combinations of verb and intensifier have fully lexicalised into fixed expressions with a new non-compositional meaning that is no longer defined by the semantics of the verb. According to Van Dale Online dictionary, *zich de tanden stuk bijten op* (lit. 'to bite himself the teeth broken on') does not mean 'to bite very hard' but something along the lines of 'to be defeated by something'. *Zich dood staren/doodstaren op* (lit. 'to stare himself dead at') and *zich blind staren/blindstaren op* (lit. 'to stare himself blind at') both mean 'to see something as the only possibility' or 'to heavily focus on something' rather than 'to stare intensely'. *Zich de ogen uitkijken* is not in Van Dale as such, but the possessive variant *zijn ogen uitkijken* 'stare his eyes out' is listed with the meaning 'to be surprised by/to marvel at the view of something'. It would of course be interesting to investigate how these expressions have lexicalised into fixed chunks, but this would take us too far afield for present purposes.

For this study, we are only interested in combinations that still have some level of semantic compositionality, meaning that the postverbal element must be interpretable

as an intensifier, boosting the original lexical semantics of the verb to a higher degree.[22] Other combinations than the ones in (48) to (51) appear to be gradually fossilising into a fully fixed lexical expression, but they have not quite reached that point yet. For instance, Van Dale mentions *het vuur uit zijn sloffen lopen* 'lit. to run the fire out of his slippers' (note that this is the possessive variant, not the fake reflexive one) as a verbal expression with the meaning 'to put in a lot of effort for something or someone'.[23] This would suggest that this combination no longer meets the criterion of semantic compositionality and is to be discarded as a fixed expression with a holistic, non-compositional meaning, much like the examples in (48) to (51). Indeed, the actual activity of running seems to be at the very least heavily backgrounded in examples like (52) below.

(52) Net zoals in liefdesrelaties **lopen** bedrijven **zich het vuur uit de sloffen** voor hun toekomstige klant zolang ze die nog moeten verleiden. (SoNaR-BE)
[...] run companies themselves the fire out of the slippers [...]
'Just like in a relationship, companies put a lot of effort into future customers when they still have to seduce them.'

At the same time, however, we do find examples in which the activity of running is still manifest – unlike the activity of biting in (48) – , and which in that regard do seem to qualify as free verb-intensifier combinations:

(53) Maar zijn ploeg van jonge sommeliers **loopt zich het vuur uit de sloffen** om de grootste wijnwonderen van de wereld te presenteren. (SoNaR-BE)
[...] runs itself the fire out of the slippers [...]
'But his team of young sommeliers is running their socks off to serve the most wondrous wines in the world.'

Moreover, there are some sporadic examples in which other verbs than *lopen* 'to run' or a different kind of footwear are selected depending on the context the phrase is used in.

(54) Jodts had een punt, want hij **reed zich** een ganse koers **het vuur uit de sloffen**. (SoNaR-BE)
[...] because he rode himself an entire race the fire out of the slippers
'Jodts had a point because he raced fiercely all race long.'

[22] Note that the exact nature of this "boosting" or "intensification" may differ slightly depending on the verb it combines with. See §3.3.6 infra for a more detailed account.

[23] A word of caution is in order, because Van Dale is not always consistent in its treatment of verb-intensifier combinations. Some combinations are listed as "uitdrukking" (expression) under the lemma of the verb or under the lemma of the intensifier, but most of the time they do not receive any label at all. Furthermore, multiple combinations of verb + *dood* 'dead' have separate entries as inherently reflexive verbs (e.g. *zich doodergeren*, *zich doodlachen*, *zich doodschrikken*...), even though these verbs are found to combine with many other intensifiers, some combinations even being more frequent than the one with *dood*, cf. Chapter 4.

(55) Arme scouts. **Lopen zich het vuur uit de sportschoenen** voor hun club, en dan zien ze dat er sprake is van erosie van clubliefde. (SoNaR-NL)
[...] run themselves the fire out of the trainers for their club [...]
'Poor scouts. They run from pillar to post for their club only to find out that the love for the club is disappearing.'

Taken together, the above examples support the idea that there is some degree of compositionality in the use of *het vuur uit de sloffen* 'the fire out of the slippers' in combination with a (limited) range of verbs. Whether the collocation with *lopen* 'to run' is indeed on the verge of lexicalisation or not is a question that we will come back to in the diachronic data investigation in Chapter 5. In this section, we merely wanted to illustrate why we have opted for the inclusion of all sentences with *het vuur uit de sloffen* (and its variants), while excluding examples like (48) to (51). There are a couple more borderline cases which have been included for similar reasons, but as these are generally not as frequent as *het vuur uit de sloffen,* they will not be separately discussed here.

A fair question to ask at this point is how to deal with examples of the *literal* fake reflexive resultative construction. In Chapter 2, it was already argued that the difference between the intensifying and literal fake reflexive resultative is not always clear-cut: without sufficient (extra- and intra-textual) context, it can be difficult to infer which meaning the speaker means to convey. As there are numerous postverbal phrases that may theoretically alternate between use as an intensifier or as a resultative phrase, we have decided to not immediately discard potential examples of the literal fake reflexive resultative construction and we will return to this in the final round of the search procedure (§3.2.1.3).

Finally, we omitted all examples that were hard to interpret for reasons of text quality and readibility. This leaves us with 3,976 relevant examples of the (intensifying) fake reflexive resultative construction. If a sentence contained two or more relevant examples of the construction (e.g. juxtaposition of multiple verbs with the same intensifier or of multiple verb-intensifier combinations), each of these was considered and counted separately and will later receive separate annotation. As many of the input intensifiers were originally found on the Internet and in other informal sources, it is hardly surprising that our query in journalistic data did not return relevant hits for a considerable number of input intensifiers. Still, the data set contained no less than 93 (potential) intensifiers, 27 of which were hapaxes, in combination with 254 different verbs (see Appendices III-2 and III-3).

3.2.1.2 Round 2: [REFL V] – [V REFL]

If we take into account the creative use of the intensifying fake reflexive resultative construction that was illustrated in Chapter 2, it is highly unlikely that the list of input intensifiers from round one represents the full extent of attested intensifiers in our

newspaper data. In this round, we aim to supplement the data set with relevant examples featuring intensifiers that were not part of the original set, by using the verbs that were culled from the corpus in the first round of the search procedure as input for a new search query. Given the large variety of verbs, we did not repeat this step for every single verb in order to limit the amount of data. Moreover, if we look at the results from the first round of our investigation, it appears that verbs which are highly frequent in the data set co-occur with a much larger variety of intensifiers than verbs with a rather low frequency. Therefore, we are arguably more likely to find new, creative intensifiers with verbs that are frequently used in this construction than with verbs that only have a few occurrences. Taking that into consideration, we only used the verbs that had at least 10 hits in round one, leaving us with the set of 35 verbs in (iii). Together, these verbs cover 3,431 or 87% of the tokens and account for 87 out of the 93 intensifiers from round one:

(iii) *bellen* 'to ring', *betalen* 'to pay', *dansen* 'to dance', *denken* 'to think', *drinken* 'to drink', *eten* 'to eat', *fietsen* 'to bike', *kopen* 'to buy', *lachen* 'to laugh', *lezen* 'to read', *lopen* 'to run', *peinzen* 'to think', *piekeren* 'to worry', *praten* 'to talk', *rennen* 'to run', *rijden* 'to drive/ride', *roken* 'to smoke', *schieten* 'to shoot', *schreeuwen* 'to scream', *schrikken* 'to be startled', *spelen* 'to play', *trainen* 'to train', *vechten* 'to fight', *vliegen* 'to fly', *vreten* 'to stuff', *werken* 'to work', *zich amuseren* 'to enjoy oneself', *zich ergeren* 'to be annoyed', *zich generen* 'to be embarrassed', *zich schamen* 'to be embarrassed', *zich vervelen* 'to be bored', *zingen* 'to sing', *zoeken* 'to search', *zweten* 'to sweat', *zuipen* 'to booze'

The corpus was then queried for all instances of these verbs – with the [lemma = "*verb*"] option in OpenSoNaR – that were followed or preceded by one of the reflexive pronouns in (i) within a range of five words. Note that in round one, we only queried the corpus for the [REFL INT] word order (i.e. the intensifier following the reflexive pronoun), whereas we are now including both the [REFL VERB] and the [VERB REFL] word orders. The intensifier query was more restricted than the verb query because the word order in which the intensifier precedes the reflexive pronoun (e.g. *Dóód*$_{[INT]}$ *schrokken we ons*$_{[REFL]}$, lit. 'Dead we startled ourselves') is extremely marked and infrequent. With respect to the verb, however, both word orders are quite common (e.g. *Hij is zich*$_{[REFL]}$ *gisteren dood geschrokken*$_{[VERB].}$ 'He startled himself dead yesterday' – *Gisteren schrok*$_{[VERB]}$ *hij zich*$_{[REFL]}$ *dood* 'Yesterday he startled himself dead'). The new search string yielded 149,081 total hits, which – after manual processing, cf. round one – were found to contain 119 new examples of the construction, featuring 40 new potential intensifiers (see Appendix III-2). However, not all of these items are really *novel* in the strict sense, because some of them bear a striking resemblance to intensifiers that were already included in the first round of the search procedure; for example *pleuris* 'pleurisy' and *te pleuris* 'to pleurisy' are of course related to *de pleuris* 'the pleurisy' and *bicblauw* 'lit. pen blue' or *donkerblauw* 'dark blue' are more specific colour variants of *blauw* 'blue'. In section 3.3.3, we will address the question

as to whether such variants should be listed as individual intensifiers or not. Furthermore, some of these lexical items (e.g. *lazerus* 'plastered' and *murw* 'mellow') were actually used as (metaphorical) resultative phrases, but we decided to include them for now because the final search round of the procedure (cf. infra) may still reveal their intensifying potential.

3.2.1.3 Round 3: [REFL INT]

The next and final step is to enter the newly discovered lexical items from round two into the same search string that was used in the first round (reflexive pronoun followed by an intensifier within a span of five words) in order to retrieve additional examples of these intensifiers with verbs that were not included in the second round, i.e. all verbs from round one that had less than ten occurrences, or any new verbs that were not yet included in our data set. The final round resulted in an additional 1,992 hits containing 14 examples of the construction, featuring 8 new verbs to add to the set of verbs from round one (see Appendix III-3).

After having gone through all three rounds of the search procedure, we re-evaluated the examples of the literal fake reflexive resultative construction. All lexical items that alternated between use as an intensifier or as a resultative phrase were retained, but the literal resultative examples received a separate label (cf. infra, §3.3.5). The reason behind this strategy is that it may be instructive to see how the intensifying fake reflexive resultative has developed in relation to its literal counterpart. However, we discarded all elements that were found to *exclusively* function as a resultative phrase in the entire data collection from the data set. By this we mean all lexical items that were found to denote some kind of result of the verbal activity in every single attestation in the data set, or for which an intensifying reading was extremely unlikely in all occurrences. This includes both instantaneous effects of the verbal activity, as well as more long-term effects that are the result of performing the verbal activity repeatedly or continuously over a certain period of time (see §3.3.5 and §3.3.6 for a more detailed description of the different types of resultative phrases versus intensifying phrases). All in all, the number of lexical items that were excluded for this reason was quite low, since the input phrases in the first round were selected on the basis of their being mentioned as potential intensifiers in the existing literature (cf. 3.2.1.1, Appendix III-1). Still, not all of these input intensifiers actually realised their intensifying potential in *our* data set. For that reason, we deleted all 8 sentences with *laveloos* 'plastered' and all 10 sentences with *rond* 'round': both adjectives were consistently used as resultative phrases, in combination with verbs of

drinking (56) and eating (57) respectively, meaning something like 'to drink/eat until you are totally plastered/full'.[24]

(56) Uitzinnige vreugde houdt onder meer in: **zich laveloos zuipen**, zijn broek afsteken... (SoNaR-BE)
[...] himself plastered to booze [...]
'Frenetic joy implies, among others: to booze until you are plastered, to pull your pants down...'

(57) Wie aan boord wil diëten, kan dat. Maar wie beslist heeft **zich rond** te **eten**, is hier eigenlijk meer op zijn plaats. (SoNaR-BE)
[...] who decided has himself round to eat [...]
'Those who want to go on a diet on board can do so. But those who have decided to stuff themselves with food, will feel more at home.'

Likewise, all 40 hits with *lazerus* 'plastered', 6 hits with *dik* 'fat' and 5 hits with *vet* 'fat' (and any combinations with these), which were retrieved in round two, were omitted because *lazerus* was always used in its "resultative" sense of 'extremely drunk', in combination with verbs of drinking, see (58), and *dik* and *vet* 'fat' retained their resultative meaning in combination with verbs of eating, see (59) and (60).[25]

(58) De betrokken cipiers **dronken zich lazerus** met binnengesmokkelde sterke drank en vernederden de gedetineerden. (SoNaR-BE)
the involved wardens drank themselves loaded [...]
'The wardens involved drank until they were loaded on smuggled liquor and they humiliated the prisoners.'

(59) Pereira is een oudere man die **zich dik eet en drinkt** aan kruidenomeletten en liters lemonade. (SoNaR-BE)
[...] who himself fat eats and drinks [...]
'Pereira is an older man who gets fat on spicy omelettes and litres of lemonade.'

(60) De meeste Amerikanen leiden een designer-bestaan, schrijft hij. Ze **eten zich vet** aan hamburgers, besteden hun geld aan overbodige rommel. (SoNaR-NL)
[...] they eat themselves fat [...]
'Most Americans lead the life of a designer, he writes. They get fat on hamburgers and spend their money on junk they don't need.'

[24] A quick Google search (performed on August 24, 2017) does give us some intensifying uses of *laveloos* and *rond*, e.g. *zich laveloos schrikken* (lit. to startle oneself plastered, 'to be very startled') and *zich rond lachen* (lit. to laugh oneself round, 'to laugh intensely'). However, as they were not found to alternate between intensifying and resultative uses in our data set, we did not include them in the investigation.

[25] Among the intensifying phrases we do find *het lazerus/lazarus*, which was listed as a fictitious disease but is etymologically related to leprosy (see WNT). Presumably, this NP intensifier was derived from the (literal) adjectival phrase.

After three rounds of search queries and manual data selection, during which we processed a total of 174,455 hits, we ended up with a data set of 4,008 relevant occurrences of the (intensifying) fake reflexive resultative construction in present-day Belgian and Netherlandic Dutch (i.e. a precision rate of 2.3%).[26] The data set contained 122 different intensifier types, 47 of which are hapaxes, and 260 verb types (see Appendices III-2 and III-3 for a list of all intensifiers and verbs). To conclude, the total verb type count, intensifier type count and intensifier hapax count in the final version of the synchronic data set can be found in Table 3.4 below.

Table 3.4. General frequency information of the final version of the synchronic data set

	VERB TYPES	INTENSIFIER TYPES	INTENSIFIER HAPAX LEGOMENA	SIZE OF DATA SET
Belgian Dutch	185	98	45	2,818
Netherlandic Dutch	152	68	23	1,190
TOTAL (MINUS OVERLAP)	260	122	47	4,008

Before we proceed to the discussion of the linguistic annotation of these sentences, the next section will go into the compilation of the diachronic data set on the basis of the Delpher corpus. So as to ensure maximal comparability, the procedure for data retrieval and selection essentially follow the same protocol and the same selection criteria but there are a number of slight modifications – due to the difference in the internal design of the two corpora – that call for some additional clarification.

3.2.2 Diachronic data set: Delpher corpus

In this investigation we will trace the recent history of the intensifying fake reflexive resultative construction in order to seek an explanation for the intriguing fact that, in present-day Dutch, the construction exhibits a high degree of productivity and creativity on the one hand, whereas on the other hand, there exist a number of preferred collocations that seem to keep this creativity within certain limits (see Chapter 2 for theoretical discussion and Chapters 4 and 5 for a detailed analysis of this mix of productivity and lexical idiosyncrasy). The present-day situation being the starting point of this investigation, the synchronic data are accordingly used as the point of departure for the compilation of the diachronic data set. As the Delpher corpus only contains Netherlandic Dutch (cf. §3.1.3), only the Netherlandic part of the synchronic data set will be used as a point of comparison for the diachronic part of the study.

[26] The precision rate measures the number of relevant instances that are left after having filtered out all false positives (i.e. hits that were retrieved by the query but which are not relevant to the investigation.)

Parallel to the cyclic search procedure that was described in the previous section, we followed a multi-step procedure to cull relevant instances from the Delpher corpus. However, the diachronic investigation will be limited to two, rather than three, search rounds. Given that the amount of extra data that was obtained in the final round during the synchronic investigation was rather limited, and in view of the much larger size of the Delpher corpus and the overall low precision rate (cf. infra), we have decided to not proceed to a third round. Another difference lies in the lexical input in the first round, which is now verb-based rather than intensifier-based. In Chapter 2, it was argued that the domain of intensification is characterised by a process of constant lexical renewal, with intensifiers emerging and falling out of use to keep pace with linguistic fashion. If this pragmatic wear-and-tear also applies to the intensifying fake reflexive resultative construction, we may expect to find drastic changes in the lexical elements that have filled the intensifier slot over the past two centuries. In that regard, it would be counterproductive to use the set of intensifiers that are found in present-day Dutch as the point of departure for the diachronic part of the investigation. Conversely, there is no reason to expect a significant change in the verbs that are intensified, as the types of verbal activities that are prone to intensification are very unlikely to display dramatic change.

3.2.2.1 Round 1: [VERB REFL] – [REFL VERB]

The synchronic data set that was constructed on the basis of the SoNaR corpus contained a total of 1,190 hits for Netherlandic Dutch, featuring 152 different verb types in the construction. Considering that the search query is once more expected to generate a substantial amount of noise and that the Delpher corpus is considerably larger than the SoNaR corpus, it would be an enormous task to take into account all 152 verbs. In order to track which lexical items have appeared in the construction as intensifiers in the 19th and 20th Centuries, we have limited the search to the top ten of intensified verbs in the synchronic data set. This means that in selecting the verbs, we only looked at the examples of the *intensifying* fake reflexive, not taking into account the literal resultative examples which may artificially inflate the frequency of certain verbs (e.g. *schieten* 'to shoot' or *rijden* 'to ride/drive'). The selected verbs are *betalen* 'to pay', *lachen* 'to laugh', *lopen* 'to run', *piekeren* 'to worry', *schrikken* 'to be startled', *werken* 'to work', *zich ergeren* 'to be annoyed', *zich schamen* 'to be embarrassed', *zich vervelen* 'to be bored' and *zoeken* 'to search'. Together, these verbs covered 766 of the 1,042 intensifying tokens and 53 of the 68 intensifier types, or respectively 74% and 78%, of the synchronic data set (NL, INT). We queried the selected decades in the Delpher corpus (cf. §3.1.3) for the following verb

forms, co-occurring with one of the reflexive pronouns in (i) within a window of five words to the left or right.[27]

(iv) *betalen, betaalen, betaelen, betaalden, betaelden, betaald, betaeld, betaelt, betaalt, betaalde, betaelde, betalend, betaalend, betaelend, betaal, betael, schrikken, schricken, geschrickt, geschrikt, schrick, schrik, schrickt, schrikt, schrikten, schrickten, schrokken, schrocken, schrikte, schrickte, schrok, schrock, geschrokken, geschrocken, schrikkend, schrickend, lopen, loopen, loopten, liepen, gelopen, geloopen, geloopt, loop, loopt, loopte, liep, lopend, loopend, lachen, lagchen, lachten, lagchten, gelachen, gelagchen, lach, lagch, lacht, lagcht, lachte, lagchte, lachend, lagchend, loech, schamen, schaamen, schaemen, schaamden, schaemden, geschaamd, geschaemd, geschaamt, geschaemt, schaam, schaem, schaamt, schaemt, schaamde, schaemde, schamend, schaamend, schaemend, vervelen, verveelen, verveelden, verveeld, verveel, verveelt, verveelde, vervelend, verveelend, ergeren, ergerden, ergert, ergerde, erger, ge*rgerd, ge*ergert, ergerend, werken, wercken, werkten, werckten, gewerkt, gewerckt, werk, werck, werkt, werckt, werkte, werckte, werkend, werckend, piekeren, pikeren, piekerden, gepiekerd, gepiekert, pieker, piekert, piekerde, piekerend, zoeken, zochten, zogten, gezocht, gezogt, zoek, zoekt, zocht, zogt, zoekend*

The performed search returned 205,537 hits over the ten selected decades. Again, we manually skimmed all the retrieved instances to identify the relevant occurrences of the (intensifying) fake reflexive resultative construction, weeding out all irrelevant hits in the process, according to the same selection criteria that were discussed in section §3.2.1.1. The Delpher database even provided additional support in favour of our decision to include the intensifying phrase *het vuur uit de sloffen* 'the fire out of the slippers', with even more examples with different verbs (61) and different kinds of footwear (62) or other objects (63).

(61) Achterhaald is het beeld van de beminnelijke oudere dame, die **zich het vuur uit de sloffen vergadert** over een onderdak voor thuisloze zwerfpoezen. (Delphcorp, 1990-1995)
[...] who herself the fire out of the slippers meets [...]
'The idea of a lovely old lady who meets with a bunch of people, trying to find shelter for stray cats, is outdated.'

(62) Als die zei: "Jongens ik reken op jullie", dan **liepen** wij **ons het vuur uit de spikes**. (Delphcorp, 1970-1979)
[...] then run we ourselves the fire out of the spikes
'When he said: "Boys I'm counting on you", we ran our hearts out.'

[27] As the Delpher corpus is not lemmatised (§3.1.3), we could not use the lemma shortcut that was employed when searching the present-day corpus through the OpenSoNaR interface. All spelling variants are based on the 19th- and 20th-Century citations in the WNT (cf. footnote 19). The wildcard in the past participle of *ergeren* 'to annoy' is added because the OCR-software had trouble with recognising diaeresis (e.g. *geërgerd* is rather consistently returned as *geÃ«rgerd*.)

(63) Terwijl de Panasonic Profwielerploeg **zich** morgen **het vuur uit de spaken rijdt** tijdens de Elfstedentocht [...] (Delphcorp, 1980-1989)
while the panasonic pro-cycling team itself tomorrow the fire out of the spokes rides [...]
'While the Panasonic Pro-Cycling team rides its heart out during the Elftstedentocht [...]'

After having filtered out all irrelevant hits, we are left with 3,171 relevant occurrences. This number still includes all examples of the literal fake reflexive resultative construction, which will be re-evaluated at the end of the search procedure (the reason behind the initial inclusion of literal fake reflexives was explained in sections §3.2.1.1 and §3.2.1.3 above). Table 3.5 gives an overview of the retrieved data for the selected decades: not only has the raw observed frequency of the construction in general gradually increased over time, so has the number of different intensifier types.

Table 3.5. Frequency results of the search query used in R1 of the construction of the diachronic data set

DECENNIA	TOTAL HITS	RELEVANT HITS AFTER ROUND 1	INTENSIFIER TYPES
1810-1819	324	0	0
1830-1839	2,604	4	3
1850-1859	5,817	12	6
1870-1879	15,439	56	14
1890-1899	16,787	71	16
1910-1919	22,899	111	25
1930-1939	28,038	227	37
1950-1959	28,404	501	57
1970-1979	33,169	672	85
1990-1995	52,056	1,517	95
TOTAL (MINUS OVERLAP)	205,537	3,171	171

3.2.2.2 Round 2: [REFL INT]

In the second search round, the Delpher corpus was queried for the 171 intensifier types from round one followed within a span of five words by the reflexive pronouns in (i). The goal of this search query is to find additional occurrences of these intensifiers in combination with verbs that were not part of the query in round one, i.e. all verbs that were not among the top ten of intensified verbs in the synchronic data set. In analogy to the first search round in the construction of the synchronic data set (§3.2.1.1), the input items were slightly modified to allow for spelling mistakes and (formal) variants (thus possibly retrieving new intensifiers), while also keeping the extra amount of noise within check (cf. supra).

In this round, another 94,078 hits were obtained, divided over the ten selected decennia as shown in Table 3.6. Once more, intensive manual post-processing was performed to sort out the relevant from the irrelevant hits, following the above-

mentioned selection criteria. As this is the last round in the search procedure, we reassessed all examples of the literal intensifying fake reflexive resultative construction and discarded all examples featuring non-alternating lexical items (i.e. lexical items that were found to exclusively function as resultative phrases in the data set, see §3.2.1.3). The lexical items that were rejected were *arm* 'poor' (37 instances), *gezond* 'healthy' (18 instances) and *in de kreukels* 'smashed up, lit. in the creases' (3 instances), the resultative uses of which are illustrated in the examples below.

(64) De door de mode geobsedeerde vrouw **koopt zich arm** aan nouveautés. (Delphcorp, 1950-1959)
[...] obsessed woman buys herself poor [...]
'The fashion-obsessed woman spends all her money on novelties.'

(65) Wie **zich gezond** wil **lagchen**, indien dat middel tot zijn herstel kan dienen, bezoeke deze salon. (Delphcorp, 1850-1859)
who himself healthy wants to laugh [...]
'Anyone who wants to laugh himself back to health, if doing so would help his recovery, should visit this saloon.'

(66) Als de trucker iets ziet en naar links moet, moet de begeleider meteen aan de bel trekken om te voorkomen dat iemand **zich in de kreukels rijdt**. (Delphcorp, 1990-1995)
[...] someone himself in the creases drives
'If the trucker sees something and has to swerve to the left, the companion immediately has to ring the bell to prevent someone from crashing.'

We also deleted all 36 instances of the NP+APs *de keel/kelen schor* and *de keel/kelen hees* 'the throat(s) hoarse', see the examples below. The reason for excluding these NP+APs is that they were always combined with verbs of yelling and thus most likely intended as literal results.

(67) Vechtende, schreeuwende fotografen, verslaggevers die **zich de keel schor schreeuwen** om een landgenoot te 'strikken'. (Delphcorp, 1970-1979)
[...] who themselves the throat hoarse scream [...]
'Fighting, screaming photographers, reporters who scream until their throats are hoarse to 'snare' a compatriot.'

(68) 35.000 toeschouwers **juichten zich de kelen hees** voor de kleine, 23-jarige ex-slager uit Luxemburg. (Delphcorp, 1950-1959)
[...] cheered themselves the throats hoarse [...]
'35,000 spectators cheered until their throats were hoarse for the little, 23-year-old ex-butcher from Luxemburg.'

Admittedly, an intensifying reading, in which case the verbal phrases in (67) and (68) should be interpreted as screaming or cheering loudly/with fervour, is not entirely impossible and the difference between the two interpretations may all in all be rather small (cf. 3.3.5 on semantic vagueness). Then again, given that a hoarse throat is a very common (short-term) result of shouting loudly, we argue that the balance ultimately tips

in favour of a resultative reading in all of the attested instances. We did decide to include the AP intensifier *schor* 'hoarse', in which the explicit mention of *keel* 'throat' is missing. Although *schor* is also primarily combined with verbs of (loud) noise emission like *schreeuwen* 'to scream', *juichen* 'to cheer' and *roepen* 'to shout', in which case a resultative interpretation is again rather likely, some examples are less obvious, see (69) and (70). As will be discussed in Chapters 4 and 5, it is plausible that intensifiers that carry less lexical content, i.e. are less lexically specific, may be less constrained by their original lexical semantics. In this case, the absence of the explicit mention of *keel* 'throat' may explain why *schor* 'hoarse' has a wider scope and is more readily interpreted as an intensifier than *de keel schor* 'the throat hoarse'.

(69) Het publiek **lachte zich zo schor** als Neelie blies, en zijzelf had dermate veel pret. (Delphcorp, 1990-1995)
the audience laughed itself so hoarse [...]
'The audience laughed so hard when Neelie blew, and she was having so much fun.'

(70) Dan gaan zij op in de Staat en verrichten net zo lang parlementaire arbeid, **praten zich schor** op vergaderingen. (Delphcorp, 1950-1959)
[...] talk themselves hoarse in meetings
'Then they are subsumed by the State and perform parliamentary tasks, they talk their heads off (all they do is talk) in meetings.'

This eventually leaves us with a total number of 5,325 relevant instances of the (intensifying) fake reflexive resultative construction, divided over the decennia as shown in Table 3.6. This number includes the hits from round one and round two, minus the overlap sentences (that is, sentences which were retrieved by both search queries because they contained one of the ten input verbs from round one). If we look at the precision rate, we find that until the mid-20th Century, the proportion of relevant to total hits does not even hit the 1% mark; the best scoring decennium is the most recent one with 3.4% precision rate. The increase in precision rate could in part be explained by a change in the quality of the corpus material. The OCR-accuracy of the Delpher database is rather unpredictable, especially for the older newspapers in the corpus, as a result of which a lot of hits had to be deleted because they were illegible or uninterpretable; in general the text recognition in recent decennia is of a much higher quality. However, an important additional explanatory factor for the observed rise in the precision rate across the decennia may be that the construction itself has undergone important changes in recent history – and this is exactly what we aim to investigate in this thesis. As expected based on the results from the first round, Table 3.6 suggests that both the number of relevant example sentences and the number of verb and intensifier types gradually increase over time. These and other aspects of frequency development, such as the question whether the construction in general has increased its overall frequency, will be addressed in more detail in Chapter 5.

Table 3.6. Frequency results of the search query used in R2 of the construction of the diachronic data set

DECENNIA	TOTAL HITS	RELEVANT HITS AFTER ROUND 2 (INCL R1)	PRECISION RATE	INTENSIFIER/ RP TYPES	VERB TYPES
1810-1819	256	0	0.00%	0	0
1830-1839	1,347	11	0.28%	6	8
1850-1859	2,674	19	0.22%	9	11
1870-1879	8,077	112	0.48%	21	36
1890-1899	10,235	196	0.73%	25	42
1910-1919	12,877	271	0.75%	31	54
1930-1939	12,135	372	0.93%	42	52
1950-1959	14,193	783	1.84%	78	67
1970-1979	12,715	1,101	2.40%	101	110
1990-1995	19,569	2,460	3.43%	118	155
TOTAL (MINUS OVERLAP)	94,078	5,325	1,79%	210	289

The next section provides a detailed explanation of the linguistic annotation of the data.

3.3 Annotation

3.3.1 Variety: Belgian Dutch or Netherlandic Dutch

The SoNaR corpus contains text material from both Belgian and Netherlandic newspapers and periodicals. It was mentioned in Chapter 2 that, aside from the salient pronunciation variation, there is ample evidence in the literature of differences in the lexicon (e.g. Geeraerts et al. 1999, Debrabandere 2005) and morphosyntactic aspects of the language (see, e.g., Haeseryn 1996 for an overview of grammatical differences, or Tummers et al. 2005 on inflectional variation), as well as in the use of certain constructions (see, e.g., Grondelaers et al. 2008 on the use of constructions with *er*, or Speelman & Geeraerts 2009 on constructions with *doen* and *laten* in Belgian and Netherlandic Dutch). What is more interesting for the current investigation is that national variation has been found within the domain of intensification as well, although the focus is more on variation in colloquial and substandard varieties. Given that intensifiers are sensitive to losing their expressive force and may be used as identity-markers, it is no surprise that regional or local dialects may use intensifiers that have not spread beyond that specific area (this is also found in other languages, cf. Chapter 2). Reker (1996), for example, reports the results of a survey aimed at making an inventory of some typical regional intensifiers (more specifically

elative compounds, which Reker calls *dikke woorden*, 'lit. fat words') in the Netherlands. Some of the more striking examples are *snotgaar* 'lit. snot-cooked' and *snotriep* 'lit. snot-ripe' for North Brabant and Limburg, *bragelvet* 'lit. mud-fat' and *strontdeurnat* 'lit. shit-soaking-wet' for Drenthe and Groningen or *breajong* 'lit. bread-young' and *dweiltrochwiet* 'lit. rag-soaking-wet' for Friesland. Both Hoppenbrouwers (1991) and Hoeksema (2012) mention the (North and South) Brabantic use of the elative element *kei-* 'rock', as in *keileuk* 'lit. rock-nice' or *keislim* 'lit. rock-smart'. De Clerck & Colleman (2013) discuss the intensifying use of *massa's* 'masses', originally a quantifier, to boost verbs, adjectives or adverbs, as in *massa's genieten* 'to enjoy a lot', *massa's slim* 'very smart' or *massa's goed gedaan* 'very well done'. Its use is restricted to the informal language of young speakers in East and West Flanders. The case of *massa's* is compared to three other similar case studies in Norde et al. (2014), which is concerned with the grammaticalisation of individual lexical items from quantifiers to degree modifiers. Limited to Netherlandic Dutch are the intensifiers *een partij* and *tig*. Although the original lexical use of *een partij* 'a part' is still dominant, its use as a degree modifier in combination with adjectives, adverbs and verbs is quite well-attested in non-standard varieties of Northern Dutch (e.g. *een partij donker* 'very dark', *een partij moeilijk doen* 'to act very difficult', *een partij stinken* 'to smell very bad'). *Tig* was originally a numeral suffix as in *dertig* 'thirty' that came to be used as an independent form expressing an undefined quantity (see Norde 2006 for a detailed account of the debonding of *tig*). As a degree modifier, it is not used in Standard Dutch, and it remains quite rare in regional varieties of Netherlandic Dutch as well, mainly occurring with comparatives (e.g. *tig beter* 'much better', *tig mooier* 'much prettier'). The intensifier *duizend* (or in its non-standard spelling *duusd/duust*) 'thousand' is used in informal varieties in several regions in Belgium and in the Netherlands. *Duizend* mainly grades adjectives and adverbs (*duizend moeilijk* 'very difficult', *duizend vaak* 'very often'), but it is also found with quantifiers (*duizend weinig* 'very little') and verbs (*duizend geslapen* 'slept very well/long'), although such uses are quite rare. Last, but definitely not least, there is an explicit reference to national variation in the construction that is at the centre of our investigation when Cappelle (2014: 273) says that "many of the intensifications are not common in Flanders (the Dutch-speaking part of Belgium) at all. This is especially true for patterns with diseases" (see also §2.2.2). As his corpus-based case study is based on Netherlandic Dutch only, he does not give any empirical data to support his claim, but it does suggest that there are some differences in the use of specific intensifier slot fillers. The detailed corpus analysis presented in Chapter 4 will show whether Cappelle's remark can be substantiated. Although not all of the above studies place equally strong emphasis on national/regional variation as an explanatory factor, they nonetheless demonstrate that there often *is* intralingual variation in the use of specific constructions and in the preferences for certain intensifiers. Based on these findings – especially the remark by Cappelle (2014) – we will investigate whether there are indeed any national differences in the use of the intensifying fake reflexive resultative construction. In order to do so, all

occurrences in the data set were tagged as either Belgian Dutch (BE) or Netherlandic Dutch (NL). The fact that Belgian Dutch is much better represented in the SoNaR corpus overall (cf. §3.1.2) is reflected in the data set, with 2,818 examples of the construction for Belgian Dutch and 1,190 examples for Netherlandic Dutch. This variable is irrelevant in the diachronic investigation, as the Delpher corpus only contains Netherlandic Dutch newspapers.

3.3.2 Verb properties

(a) **Lemma**. Table 3.6 showed that the variety of verbs found in the intensifying fake reflexive resultative construction has gradually increased over time, with over 250 different verb types in present-day Dutch. We are interested to see which (kind of) verbs were already used in this construction very early on, and which verbs were subsequently attracted to the construction. Moreover, we want to inspect which specific verbs tend to co-occur with which intensifiers, and how these co-occurrence patterns may have changed over the past two centuries. In order to follow the expansion of the verb slot on the one hand, and the changes in collocational preferences on the other, we annotated for all individual verb lemmata. We opted not to annotate for other verbal properties, like tense or aspect, because these distinctions are not immediately relevant to the aims of this investigation. Moreover, it is likely that some of these properties, like verb tense, are primarily determined by characteristics of the genre, rather than reflecting inherent properties of the construction.

(b) **Reflexivity.** The majority of verbs that are used in the V-slot of the construction are non-reflexive verbs, i.e. verbs that generally do not occur with a reflexive pronoun – hence the term "fake" –, but that may be coerced into a reflexive pattern. This is the case for verbs like *lachen* 'to laugh', *lopen* 'to run', *schrikken* 'to be startled', *werken* 'to work', *zoeken* 'to search', etc. However, there is a relatively small group of verbs in Dutch which are inherently reflexive in the sense that they always occur with a reflexive pronoun, e.g. *zich amuseren* 'to be amused', *zich ergeren* 'to be annoyed', *zich schamen* 'to be embarrassed', *zich vervelen* 'to be bored'.[28] If such inherently reflexive verbs are used in the construction, the reflexive is not truly "fake": there is already a reflexive pronoun selected by the verb, which fuses with the REFL-slot at the constructional level. This difference has not been picked up in the (scant) literature

[28] Some of these verbs do have a non-reflexive counterpart, but these have a slightly different meaning, e.g. *zich ergeren* 'to be annoyed' versus *ergeren* 'annoy (someone else)' or *zich vervelen* 'to be bored' versus *vervelen* 'to bore (someone else)'.

on this construction, so it is unclear at this point to what extent reflexive and non-reflexive verbs will differ with respect to their development over time or their use in present-day Dutch. As it is possible that there are subtle or more pronounced differences between the reflexive and non-reflexive verbs, they received separate annotation. Figure 3.1 gives the proportion of reflexive to non-reflexive verbs in the SoNaR corpus, Figure 3.2 shows how this proportion has evolved throughout the selected decennia in the Delpher corpus.

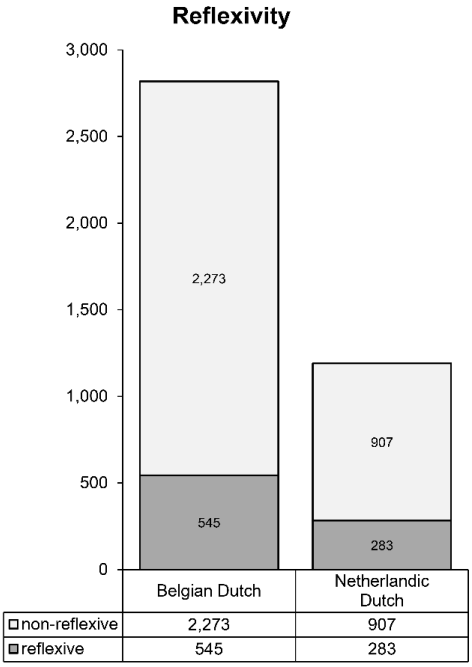


Figure 3.1. Proportion of reflexive vs. non-reflexive verbs in SoNaR

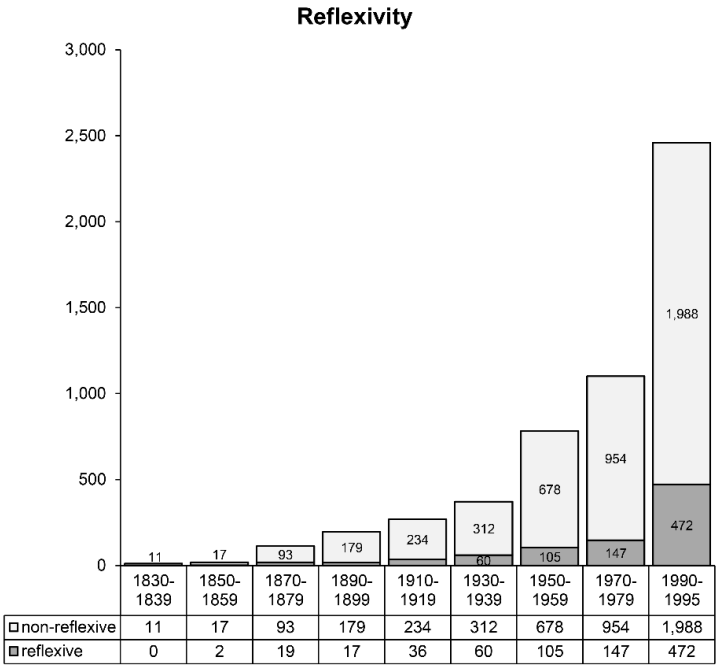


Figure 3.2. Proportion of reflexive vs. non-reflexive verbs in Delphcorp

(c) **Transitivity**. Besides the reflexivity of the verb, we also annotated all items for the transitivity of the verb. While Goldberg & Jackendoff (2004) distinguish three levels of transitivity in their work on the resultative construction, viz. intransitive, unselected transitive and selected transitive (cf. Ch2, §2.2.1), we opted for a more traditional distinction between intransitive verbs and transitive verbs. The reason for this is that the reflexive pronoun is not really a reflexive *object* selected by the verb but an inherent part of the construction, rendering the division between selected transitives and unselected transitives irrelevant. The intransitive verbs in this investigation are verbs that generally do not select for a direct object, such as *lachen* 'to laugh', *schrikken* 'to be startled', *werken* 'to work', the transitive verbs are verbs that do select for a direct object outside of this construction, such as *drinken* 'to drink', *eten* 'to eat', *zoeken* 'to search'. These transitive verbs are actually pseudo-transitive because, in this construction, they abandon their canonical object in favour of the obligatory reflexive pronoun: for example, in *Hij heeft zich een ongeluk gezocht* 'He has searched himself an accident', the object that the subject is searching for is not expressed. Figure 3.3 and Figure 3.4 give the distribution of transitive and intransitive verbs in the SoNaR data set and Delphcorp, respectively. Note that transitivity partly overlaps with reflexivity in that inherently reflexive verbs are necessarily intransitive.

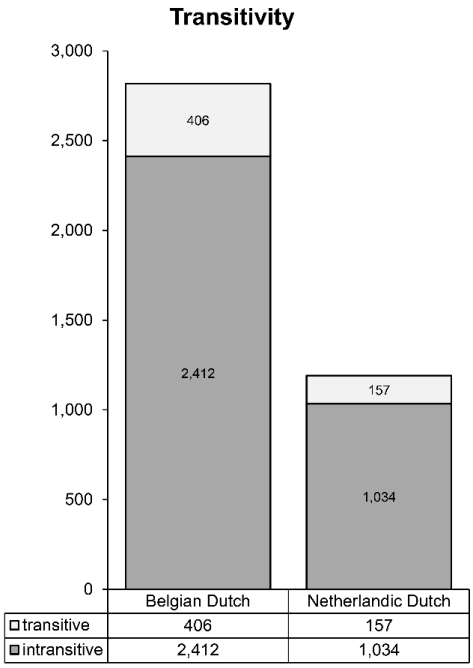
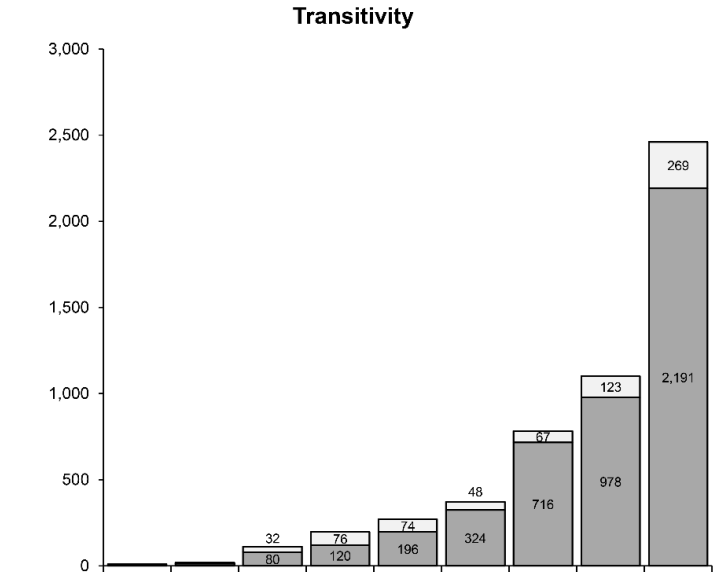


Figure 3.3. Proportion of intransitive vs. transitive verbs in SoNaR



	1830-1839	1850-1859	1870-1879	1890-1899	1910-1919	1930-1939	1950-1959	1970-1979	1990-1995
transitive	3	6	32	76	74	48	67	123	269
intransitive	8	13	80	120	196	324	716	978	2,191

Figure 3.4. Proportion of intransitive vs. transitive verbs in Delphcorp

3.3.3 Intensifier properties

(a) **Surface form and lemma**. Each item in the data set was annotated for both the surface form of the intensifier (i.e. exactly as it is represented in the sentence) and for the lemma of the intensifier, which generalises over certain spelling or morphological variations. We also noted in §3.2.1.2 that some lexical elements closely resemble one another – to the point where one is likely to have been derived from the other –, but these were still counted as individual intensifiers. The decision as to which variations are important enough to be treated as separate lemmas and which are merely spelling mistakes or inconsequential formal variants, is not a trivial one. In this section we discuss some specific special cases of this kind to motivate why we have decided to either opt for separate entries or to merge surface forms into one lemma. All decisions are summarised at the end in Table 3.7.

First of all, spelling variants are subsumed under one lemma for which the most frequent spelling is chosen as label. For example, we find vowel variants in *pleuris* versus *pleures*, *schompes* versus *schompus* and *lazerus* versus *lazarus,* consonant variants in *tyfus* versus *typhus* and variation in the presence of the linking consonant in *apezuur* versus *apenzuur*. One exception is *leplazerus* and *laplazerus*, which we treated as two individual intensifiers because both forms are more or less equally frequent and we are unsure whether the *e/a*-variation is just a random spelling feature or whether they are intended as two separate fictitious diseases. Compounds with *half* 'half' can be written in one word or in two words (e.g. *half dood* or *halfdood* 'half dead'), but are

counted as one lemma.

Second, the NP+PP category in particular presents a challenge because it allows for specific types of variation that are not attested in other categories. A number of intensifiers from the NP+PP category display some variation in the preposition and in the determiner. In addition to *de benen uit het lijf*, for instance, we find *de benen van het lijf*, *de benen vanonder het lijf*, *de benen onder het lijf uit* 'the legs out of the body', etc. but also *de benen van zijn/haar/ons lijf* 'the legs out of his/her/our body' with a possessive pronoun instead of a definite article. Another example is *de longen uit het lijf* 'the lungs out of the body', which can also be realised as *de longen uit zijn/haar/mijn lijf* 'the lungs out of his/her/my body'. The inclusion of these variants as separate types would inflate the type frequency and the hapax count, which may lead us to overestimate the productivity. In the synchronic data set, there are 6 extra variants, 3 of which are hapaxes, for de *benen uit het lijf*; for *de longen uit het lijf*, we find 3 extra variants, all of which are hapaxes. If we were to include all the variants as individual types for just these two intensifiers, our total type frequency would go from 122 to 131 and the hapax count would be raised from 47 to 53. We argue that such small formal variations do not reflect true type expansion of the construction and do not really showcase any real creativity on the part of the speaker. Therefore, variants of this nature should not be allowed to influence the productivity measures and were taken together as one lemma.

Another type of variation in the NP+PP category is the modification of the noun, either in the NP or in the PP. Some examples: *de lange benen uit het lijf* 'the long legs out of the body', *de benen uit het dampende lijf* 'the legs out of the steaming body', *de benen uit het nog jonge lijf* 'the legs out of the body that is still young', *de longen uit het tengere lijf* 'the lungs out of the frail body', *de longen uit het lijfje* 'the lungs out of the little body', *de blaren aan de edelachtbare voeten* 'the blisters on the noble feet', *de ogen uit de vaak nog baardeloze koppies* 'the eyes out of the often still beardless little heads', *een flink stuk in zijn kraag* 'a big piece in his collar'... The fact that we can still modify the nouns in the NP- or PP-part of the intensifier by means of an adjective or diminutive provides evidence that these intensifiers are not fully fixed yet, which is an important observation in itself. We also acknowledge that some of these variants may add some extra effect of expressivity to the utterance and, in that regard, do reflect the speaker's linguistic creativity. However, like the aforementioned formal variants, they do not really represent truly new intensifier types and should not be taken into account as such in the frequency analyses.

Additionally, the data set contains multiple examples of NP+PP intensifiers that either share the NP-part or the PP-part. The PP *uit het lijf* 'out of the body' recurs in *de longen uit het lijf* 'the lungs out of the body', *de benen uit het lijf* 'the legs out of the body', *de ziel uit het lijf* 'the soul of the body', *de pleuris uit het lijf* 'the pleurisy out of the body' and *de naad uit het lijf* 'the seam out of the body'. If we look at the NPs, we see that, for

example, *de ziel* 'the soul' can be complemented by *uit het lijf* 'out of the body', *uit de naad* 'out of the seam' or *uit de raap* 'out of the head' and we can get *het vuur* 'the fire' out of *de sloffen* 'the slippers', *de slofkes* 'the little slippers', *de schoenen* 'the shoes', *de sportschoenen* 'the trainers', *de molières* 'the lace-ups', etc. All of these examples have been coded as separate types because they may reveal the way in which the construction attracts new intensifier types and expands its use over time. As we will argue in Chapters 4 and 5, at least some of these appear to be examples of analogical extensions or analogisation, a process by which new intensifiers are created on the basis of (frequent) model intensifiers. This process has been found to play a crucial role in the expansion of constructions and language change in general (De Smet 2013, Traugott & Trousdale 2013, Trousdale 2014, De Smet & Fischer 2017, Norde & Strik 2017).

Finally, there are some other types of variation in other categories that need to be dealt with in short. We sometimes find the same head noun in different configurations or in different syntactic categories. To give but two examples, the word *pleuris* 'pleurisy' has three formal realisations, the NP *de/het pleuris,* the PP *te pleuris* and the AP *pleuris*[29] (cf. §3.2.1.2) and for *blubber* 'blubber', we find *de/het blubber* (singular NP), *de blubbers* (plural NP) and *te blubber* (PP). From a diachronic perspective, it is likely that one of these variants was introduced first and later came to serve as a model for the other, derived forms. If we were to conflate these into one lemma, we might do injustice to this diachronic development, which may also be motivated in part by analogical thinking on the basis of existing models (cf. supra). We have also opted for separate lemmata in compounds, like *donkerblauw* 'dark blue', *bic-blauw* 'pen blue', *grasgroen* 'grass green' or *steendood* 'stone-dead'. A conceivable explanation for the use of such compounds is that the simple intensifiers *blauw* 'blue', *groen* 'green' and *dood* 'dead' have shed some of their expressive force and are no longer felt to be sufficiently "extravagant" in specific contexts (see Ch2, §2.3, on the role of expressivity in language change). If that is true, we expect to find such compounds in more recent decennia, after *blauw*, *groen* and *dood* have reached a certain frequency. A similar motivation may explain the use of both *een slag in de rondte* 'a/one punch around' and multiple *slagen in de rondte* 'X punches around'.

[29] We could also consider *pleuris* as a noun without determiner, but given the parallel between sentences with an obvious AP, like *Hij schrikt zich rot/wild/kapot/dood* (lit. He startles himself rotten/wild/broken/dead, 'He is very startled') and *Hij schrikt zich pleuris*, we have treated *pleuris* in the latter sentence as an AP.

Table 3.7. Summary of the intensifier lemmatisation

	Separate type/lemma	Merged
Spelling variants		♦ (excl. *leplazerus/laplazerus*)
NP+PP: preposition variation		♦
NP+PP: determiner variation		♦
NP+PP: shared NP/PP	♦	
Same word in multiple syntactic categories (e.g. *pleuris*)	♦	
Compounds	♦	
Slag/slagen in de rondte	♦	

(b) **Syntactic category**. It has been repeatedly stated that the intensifying fake reflexive resultative construction displays a high degree of productivity in present-day Dutch and that this productivity testifies to the great linguistic creativity of speakers of Dutch. This is confirmed by the large variety of different lexical items that can fill the INT-slot of the construction, with a total of 122 different intensifiers in the synchronic data set and over 200 intensifiers that have been used at some point in the past two centuries (cf. §3.2.1.3 and §3.2.2.2 above). These lexical items can be recruited from multiple syntactic categories, viz. AP, NP, PP, NP+AP, NP+PP and NP+particle, but not all of these categories are equally well-represented. Some syntactic categories may, for instance, contain more frequent intensifiers than other categories. The distribution across the different syntactic categories that are represented in the synchronic data set is shown in Figure 3.5.

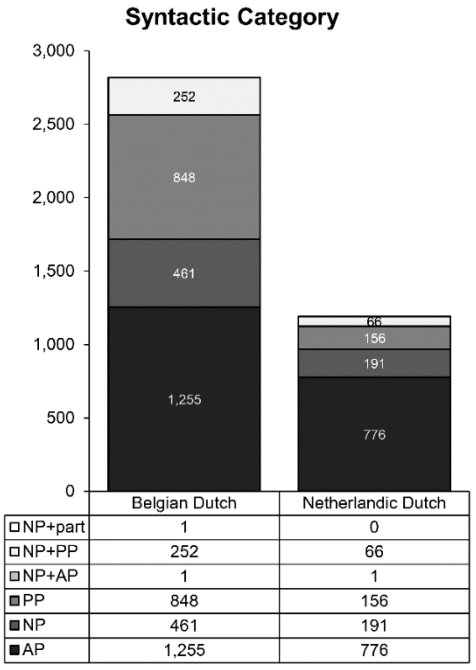
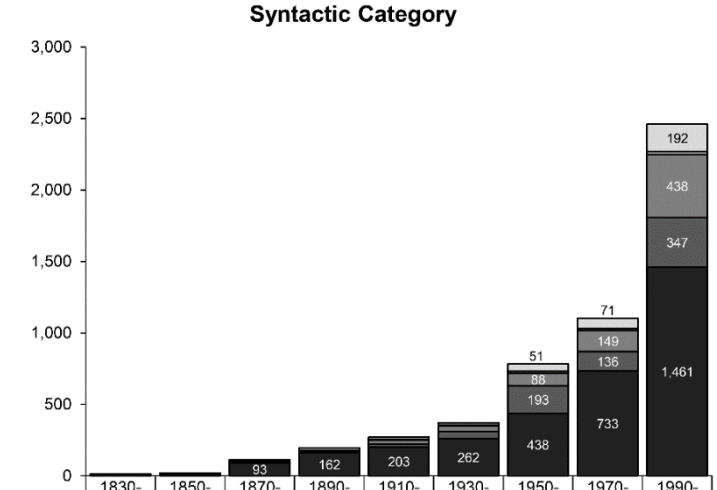


Figure 3.5. Proportion of the six syntactic categories of intensifiers in SoNaR

Table 3.6 showed that the number of different intensifiers that are found in the construction has gradually increased over time. In addition to investigating which specific intensifiers were among the first to be used in this construction in the early 19th Century and which items were subsequently added to the repertoire, we also want to see how the different syntactic categories have contributed to the repertoire of intensifiers over time. There are indications in the existing literature that some syntactic categories are more susceptible to adopting an intensifying meaning than others. There are many studies on adjectives developing into degree modifiers (Ito & Tagliamonte 2003, Vandewinkel & Davidse 2008, Lorenz 2002, Tagliamonte 2008, Margerie 2014, Wharton 2016, among others), but other word classes, like content nouns or quantifiers, have been shown to develop degree or intensifier meanings as well (see, e.g., Doetjes 2008, Norde et al. 2014, Norde & Van Goethem 2014). Figure 3.6 shows how the proportion of the syntactic categories has developed over the past two centuries. For now we are only interested in how the syntactic categories were divided over the different subsets in our corpus in terms of total frequency, but it will also be interesting to see how many different intensifier types each syntactic category contributes to the total pool of intensifiers and how this may have changed over time. This will be discussed as part of the description of the general frequency developments in Chapter 5.



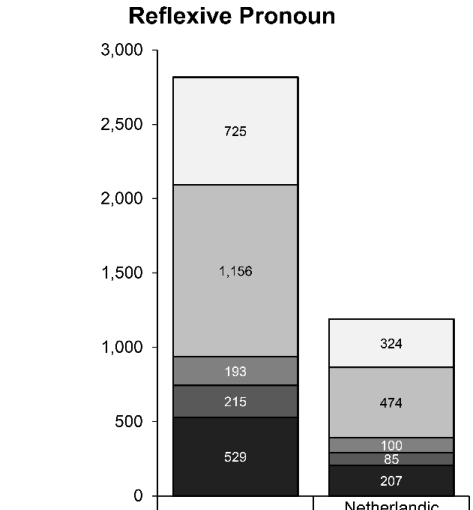
	1830-1839	1850-1859	1870-1879	1890-1899	1910-1919	1930-1939	1950-1959	1970-1979	1990-1995
NP+part	0	0	0	0	0	0	0	0	0
NP+PP	1	3	2	17	16	16	51	71	192
NP+AP	0	0	5	6	5	5	13	12	22
PP	1	3	8	9	28	42	88	149	438
NP	0	1	4	2	18	47	193	136	347
AP	9	12	93	162	203	262	438	733	1,461

Figure 3.6. Proportion of the six syntactic categories of intensifiers in Delphcorp

3.3.4 Reflexive pronoun

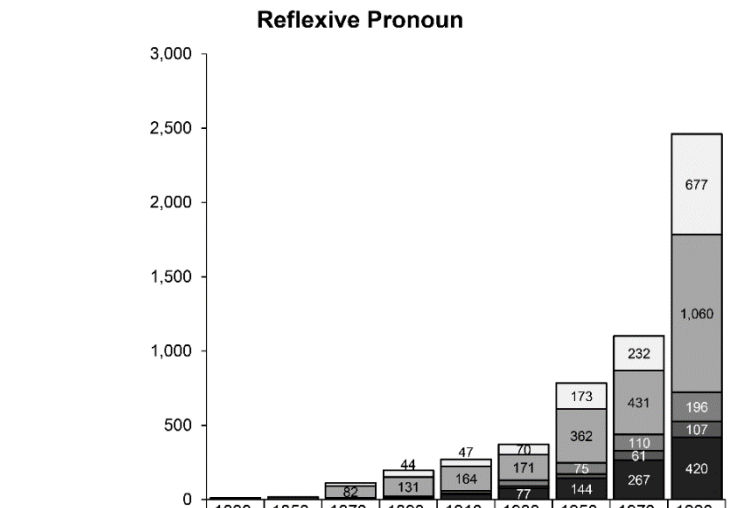
In the analyses in the following chapters, the emphasis is mainly on the realisation of the verb and the intensifier slots of the construction, because, in their status as "open" slots,

they are the ideal candidates for investigating changes in productivity. The reflexive pronoun is much less interesting in that regard because it is lexically specified and allows for very little variation. The reason for including this variable is to find out whether speakers of Dutch mainly use the construction to talk about their own feelings and activities (first and second person) or to talk about events that happened to other people. Although it is not the primary goal of our investigation, this may shed some light on the motivation behind using this construction. In turn, it may give us more insight into how this construction compares to other means of intensification or expressive language and how it fits into the domain of linguistic expressivity in general. We distinguished between the three persons (first person, second person, third person) and the two numbers (singular, plural), as shown in Figure 3.7 for the SoNaR data set and Figure 3.8 for Delphcorp. There are some finer distinctions that were not taken into account. For example, we did not make a distinction between weak and strong forms of the pronoun (e.g. *me* versus *mij*, *je* versus *jou*) because the strong forms were very infrequent overall, accounting for less than 1% of all data in both data sets. Also infrequent were emphatic pronouns (e.g. *mezelf*, *jezelf* and *onszelf* together were found in less than 0.9% and 0.3% of all instances in SoNaR and Delphcorp respectively); the only emphatic pronoun occurring with some frequency was *zichzelf* (although it occurs in only 2% of the SoNaR data and 1% of the Delphcorp data), but this was mainly caused by the fact that *zichzelf dood schieten/rijden* 'to shoot/drive oneself dead/to death' are frequently occurring resultative expressions.



	Belgian Dutch	Netherlandic Dutch
3rd pers. plural	725	324
3rd pers. singular	1,156	474
2nd pers. sing/pl	193	100
1st pers. plural	215	85
1st pers. singular	529	207

Figure 3.7. Proportion of the forms of the reflexive pronoun in SoNaR



	1830-1839	1850-1859	1870-1879	1890-1899	1910-1919	1930-1939	1950-1959	1970-1979	1990-1995
3rd person plural	2	3	19	44	47	70	173	232	677
3rd person singular	8	12	82	131	164	171	362	431	1,060
2nd pers. singular/plural	0	2	1	6	17	40	75	110	196
1st pers. plural	0	0	1	3	5	14	29	61	107
1st pers. singular	1	2	9	12	37	77	144	267	420

Figure 3.8. Proportion of the forms of the reflexive pronoun in Delphcorp

3.3.5 Literal versus intensifying semantics

In addition to the slot-filler-specific variables, we annotated all examples in our data set for the semantics of the verb-intensifier combination, thus moving beyond the lexical semantics of the verb and the postverbal phrase as viewed in isolation. We evaluated all individual items in the data set to decide whether the verb in combination with its postverbal phrase in that particular utterance should be interpreted as an example of the literal fake reflexive resultative construction (verb + resultative phrase, marked as RES) or as an example of the intensifying fake reflexive resultative construction (verb + intensifier, marked as INT).

In her study on *to death* as a degree modifier, Margerie (2011) makes a more fine-grained semantic distinction between five semantic categories, viz. degree modifier (i.e. our category of "intensifier"), hyperbolic potential result, potential result, hyperbolic actual result and actual result, though the distinction between some of these categories is not always fully clear from the examples she provides:

(71) Semantic categorisation of *to death* (Margerie 2011)

a. actual result: All Israel stoned him to death. (Margerie 2011: 121)

b. hyperbolic actual result: I'm sure she's working everybody to death. (Margerie 2011: 133)

c. potential result: My master Kasim is sick well nigh unto death. For many days he hath nor spoken nor tasted aught of food. (Margerie 2011: 125)

d. hyperbolic potential result: For I was faint and weary, and sick almost unto death. (Margerie 2011: 137)
e. degree modifier: It brings a consumed long string of past transactions, that bore me to death. (Margerie 2011: 127)

Her corpus study is aimed at elucidating how one individual item, viz. *to death*, has developed its booster function in various constructions (viz. [NP_1 V NP_2 *to death*]/[NP BE ADJ *to death*]) from the 16th Century onwards. While the question as to how individual lexical items developed a new use as degree modifier is definitely an interesting topic for Diachronic Construction Grammar, we are not focused on tracking such semantic developments for all of the intensifier phrases encountered in the data. Instead, we are taking a more general perspective by focusing on the use and variation of the schematic intensifying fake reflexive resultative pattern (i.e. [SUBJ V REFL INT]). It is remarkable, in this regard, that many of the intensifiers that are found to occur in the INT-slot do not even have a degree modifier meaning outside of this construction, as was already mentioned in Chapter 2 (§2.3). The semantic categories illustrated in (71) are mainly of importance in the so-called transitional stages and bridging contexts, because they appear to have paved the way for a degree modifier reading to arise in the case of *to death* (Margerie 2011: 139); of course, in other individual cases, quite *different* transitional stages may have been at stake. As was argued in Chapter 2, we do not wish to focus on the origins of the intensifying fake reflexive resultative construction, but we are primarily interested in the changes it has undergone after it was established as a full-fledged construction. Even so, we cannot ignore the obvious diachronic and synchronic relationship between the intensifying construction and the literal resultative construction. Not only is it likely that the intensifying construction has developed from the resultative construction via reanalysis, they are still linked by virtue of their formal pattern in present-day Dutch. As the frequency data and analyses presented in Chapters 4 and 5 will be limited to the instances that are unequivocally intensifying in meaning, the main aim of the present paragraph is to provide working criteria for delineating what falls (and does not fall) under the intensifying fake reflexive resultative construction for the purposes of the present investigation. In general, we have opted for a coarse distinction between, on the one hand, a resultative (i.e. non-intensifying) reading, in which the above categories (a) to (d) are merged, and, on the other hand a purely intensifying reading. However, although the practice of corpus research requires such a strict categorisation, it does not mean that there *is* a clear dichotomy between the two constructions and we can expect to find examples that resist straightforward classification. In this section, we will explain what guided us in classifying individual corpus instances as representing either the resultative or the intensifying reading and we will discuss some transitional uses on the basis of data from the synchronic and diachronic corpora.

For the large majority of the sentences culled from the corpora via the search procedures described earlier in this chapter, their semantic analysis as "intensifying" or

not was unproblematic. Given the intensifier-oriented bias in the search procedure (cf. supra), most postverbal phrases included in the investigation almost exclusively function as an intensifier and do not trigger potential ambiguity with a resultative interpretation. For example, our experience and world knowledge tell us that people do not actually turn purple when they are annoyed (72), or worse, get pneumonia as a result of studying (73), nor can we imagine a situation in which one receives or turns into a monkey when startled (74).

(72) Wilders heeft **zich paars geërgerd** aan het optreden van koningin Beatrix. (SoNaR-BE)
wilders has himself purple annoyed [...]
'Wilders was very annoyed by the appearance of queen Beatrix.'

(73) Jan Verheyen heeft **zich** toen **de pleuris gestudeerd** op lijstjes. (SoNaR-BE)
jan verheyen has himself then the pleurisy studied [...].
'Jan Verheyen has learnt a lot of lists by heart back then.'

(74) Die begonnen in de lucht te schieten. Ik **schrok me** natuurlijk **een aap** en riep: dekken! (Delphcorp, 1950-1959)
[...] I startle myself of course a monkey [...]
'They began shooting in the air. I was of course very startled and yelled: take cover!'

In total, there are only 14 out of the 122 intensifiers attested in the SoNaR data set and 37 out of the 210 intensifiers attested in Delphcorp that were found to function both as intensifier and as a resultative phrase in the data (see Appendix III-5). If we look into the behaviour of these potentially ambiguous intensifiers in more detail, we see that the collocational patterns in their intensifying uses are for the most part very different from their uses as a resultative phrase. There are several verb-intensifier combinations which in practice clearly only allow for a resultative interpretation, as in (75) and (76) below.

(75) Maar toen trad de directeur naar voren om **zich dood te schieten**. (Delphcorp, 1890-1899)
but then stepped the principal to front to himself dead to shoot
'But then the principal stepped forward to shoot himself dead.'

(76) Graf raakte in een bocht van het circuit in de slip en **reed zich te pletter** tegen een rots. (Delphcorp, 1970-1979)
[...] and drove himself to smithereens against a rock
'While taking a turn, Graf got off course into the mud and crashed into a rock.'

In the examples above, the resultative phrases *dood* 'dead' and *te pletter* 'to smithereens' denote relatively instantaneous effects of the verbal activities shooting and driving, respectively. Verbs can also denote activities that have non-instantaneous, *long*-term results. In example (77) below, the death of the plants is construed as an (in this case desired) end result of their stimulated rapid growth and in (78), the parents being sick is a long-term result of their having worked too hard for too long. In (79), the most plausible interpretation is that the subject hurt his hand by hitting the metal frame with a hammer

(repeatedly or over an extended period of time). Such instances are straightforward examples of the literal resultative construction, too.

(77) Van Ooijen tipt bestrijdingsmiddelen met een groeistof, die de planten van wortel tot stengel opnemen. Daardoor **groeien** ze **zichzelf dood**. (SoNaR-NL)
[...] grow they themselves dead
'Van Ooijen adds some growth substance to the pesticides, which the plants absorb from the root to the stalk. In doing so, they grow and grow until they die.'

(78) Ik geef ze geen ongelijk, mijn ouders zijn inmiddels afgekeurd... **zich ziek gewerkt.** (SoNaR-NL)
[...] themselves sick worked
'I'm not saying they're wrong, my parents have been declared unfit because they worked so hard (for so long) that they got sick.'

(79) Met een hamer **sloeg** hij **zich de handen stuk** op het metalen frame. (Delphcorp, 1990-1995)
[...] hit he himself the hands broken [...]
'With a hammer he hit the metal frame until his hands were sore.'

There are also some items for which the resultative interpretation is perhaps less straightforward, such as (80) and (81) below. These are unlikely to be interpreted as results in the actual sense of the word: in (80) the subject does not possess the gift of actually dropping dead as a result of the verbal activity (unlike in (75)) and neither did the cyclers literally fall to smithereens or crash into something in (81) (unlike in (76)). At the same time, these are not instances with a purely intensifying meaning either: paraphrases like 'she ran intensely/extensively' or 'they had cycled intensely' do not seem to adequately capture the meanings that are meant to be conveyed in the respective contexts. The verbal activity clearly brings about some kind of resultative effect, but this effect is not to be interpreted *literally* as dropping dead or crashing to smithereens. Instead, we propose that these examples illustrate hyperbolic mapping: the subjects in both sentences are so exhausted as a result of the verbal activity that it almost feels as if they could die or have crashed into something.

(80) Ze heeft de gave om **zich** helemaal **dood** te **lopen**, zoals ze zelf zegt. Ze kan hard zijn voor zichzelf. (Delphcorp, 1990-1995)
she has the gift to herself totally dead to run [...]
'She has a gift to fully exhaust herself when she is running, as she so claims. She can be hard on herself.'

(81) We hadden **ons** al **te pletter gereden** achter twee gemiste vluchten. (SoNaR-BE)
we had ourselves already to smithereens cycled after two failed breakaways
'We had already exhausted ourselves in two failed breakaways.'

If we were to follow Margerie's (2011) semantic classification above, the sentences in (80) and (81) could be seen as instantiating one of the transitional constructions expressing

hyperbolic (potential or actual) result. As our primary focus lies with internal changes in the *intensifying* fake reflexive resultative construction, such transitional constructions have been subsumed under the broad resultative (i.e. non-intensifying) category in the present investigation. Still, the examples in (80) and (81) show that the difference between the resultative and intensifying fake reflexive resultative construction is not always clear-cut. The next paragraph will address this often fine line between both interpretations by referring to theoretical notions like semantic ambiguity and vagueness. Although it may not always be possible (or necessary) to choose between one or the other, we will illustrate how contextual clues may often guide us towards the most likely interpretation.

3.3.5.1 Classification: ambiguity, vagueness and contextual clues

Because of the synchronic and diachronic relatedness of the two constructions, there are a number of cases in which both interpretations seem to flow into one another. For practical purposes, we had to draw a boundary somewhere: if we want to provide a detailed description of the use and development of the intensifying fake reflexive resultative construction, we need to be able to delineate the intensifying construction from its resultative counterpart. Some cases that we will discuss in this section are not straightforward examples of either of the two categories, but, even in such cases, there are still a number of clues that can guide us in the classification. First of all, some instances include explicit means of disambiguation. Language users may, for instance, want to signal that what they are about to say is not to be interpreted literally by adding quotation marks or modifiers such as *als het ware* 'as it were', *omzeggens* 'so to speak', *bij wijze van spreken* 'so to speak' etc.[30]

(82) Ikzelf **solliciteer me** *bij wijze van spreken* **dood**. Al anderhalfjaar. Steeds word ik om onbenullige redenen afgewezen. (Delphcorp, 1990-1995)
I-self apply myself so to speak dead [...]
'I am applying for jobs like crazy, so to speak. For over 1.5 years now. I keep getting rejected for trivial reasons.'

(83) Investeerders **lopen zich** *als het ware* **het vuur uit de sloffen** om op de Franse beurs tegen elke prijs aandelen te kunnen kopen. (SoNaR-BE)
investors run themselves as it were the fire out of the slippers [...]
'Investors are running their socks off, as it were, to buy stocks at any price on the French stock market.'

[30] It would be interesting to investigate how frequently language users make use of such explicit markers. This was not possible in the present investigation because much of the original typography (including quotation marks) was lost during the OCR process.

Or, conversely, the speaker may use the word *letterlijk* 'literally' to inform the reader that the sentence *is* to be interpreted as a literal resultative phrase, as in (84) and (85) below.

(84) In 'Gebrek aan bewijs' van Yoh Sano **schrikt** een man **zich** *letterlijk* **dood**, als tijdens het maken van een groepsfoto de fotograaf plotseling een apemasker opzet. De overleden man had een zwak hart. (Delphcorp, 1980-1989)
[...] startles a man himself literally dead [...]
'In 'Lack of evidence' by Yoh Sano, a man was so startled when the photographer put on an ape mask that he literally died. The deceased man had a weak heart.'

(85) Zelfmoord gepleegd, of **verveelde** hij **zich** *letterlijk* **dood** bij zijn vrouw? (SoNaR-BE)
[...] he himself literally dead [...]
'Did he commit suicide, or was he so bored with his wife that he actually died because of it?'

The examples with *letterlijk* 'literally' suggest that the speaker feels the need to explicitly block an intensifying interpretation to avoid misunderstanding on the part of the hearer (cf. infra on the "default" status of the intensifying meaning). However, some caution is warranted here. In examples (84) and (85), we know that *letterlijk* is indeed used in its basic literal sense because there are other elements in the sentence (*overleden man* 'deceased man' and *zelfmoord* 'suicide') that confirm that someone actually died. However, the use of *letterlijk* alone is certainly not always a reliable contextual clue for a resultative use. In the examples (86) and (87) below, the employees most likely did not get pneumonia during their work and the politicians did not turn blue as a result of their annoyance, even though it is said that they *literally* did.

(86) Ook wel bij andere diensten echter moet men **zich** *letterlijk* **de pleuris werken** door gebrek aan mankracht! (Delphcorp, 1990-1995).
[...] must one himself literally the pleurisy work [...]
'In other departments as well, people need to work very hard due to the lack of manpower.'

(87) De liberalen **ergeren zich** *letterlijk* **blauw** aan hoe bepaalde instellingen zelf bepalen welk beleid ze voeren. (SoNaR-BE)
the liberals annoy themselves literally blue [...]
'The liberals are very annoyed by the fact that some institutions themselves decide which policy to pursue.'

Once more, we find that speakers appear to have a universal need for expressivity: in the examples above, *letterlijk* is not meant to signal that the utterance is to be interpreted in a literal sense; quite the opposite, it may serve to further boost or strengthen the verb-intensifier phrase, much like the adverbs *absolutely* or *totally*. In present-day Dutch, we often find similar examples of emphatic uses of *letterlijk* 'literally' in other constructions, as in *Ik ga letterlijk dood van de stank* 'I am literally dying because of the smell' or *Het duurde*

letterlijk eeuwen voor ik mijn eten kreeg 'It literally took ages for me to get my food'.[31] On a side note, speakers with some knowledge of the use of colour in politics will understand that a pun is intended in (87), as the liberal party is often associated with the colour blue in many countries, including Belgium. The careful selection of particular intensifiers in certain contexts in order to create a humorous effect is not exceptional in this construction, see also e.g. (72) for the association of Geert Wilders and *paars* 'purple' in the context of the "purple coalition" in the Netherlands. We will return to the use of context-specific intensifiers in Chapters 4 and 5. In the absence of such explicit clues, the larger textual context, in combination with world knowledge and basic common sense generally provide the necessary information for deciding between an intensifying and a literal meaning, as was mentioned earlier.

However, there are a number of cases in which neither general world knowledge nor the linguistic context provide immediate answers. Consider the following examples:

(88) Hij **werkte zich in het zweet**, was des nachts om een uur of twee klaar. (Delphcorp, 1990-1995)
he worked himself in the sweat [...]
'He worked hard/until he was sweating, he finished around 2 AM.'

(89) Ze storten zich op het hapjesbuffet ter ere van de burgemeester, **zingen zich schor** in de kapel. (SoNaR-NL)
[...] sing themselves hoarse in the chapel.
'They threw themselves upon the buffet of appetisers in honour of the mayor, sang with fervour/until they were hoarse in the chapel.'

In these specific cases, the two readings are closely related in that working with a certain intensity may indeed cause someone to sweat and singing with fervour may induce hoarseness. In fact, both readings may be simultaneously applicable, as there is not really much of a difference in the overall interpretation of the sentence – unlike in the case of, e.g., *Hij werkt zich dood* 'he works himself dead/to death', where there is intuitively a big difference between the literal resultative and the intensifying interpretations (either one is dead or one is not). In other words, we may raise the question as to whether we really can (and have to) decide between the two readings at all in such cases. Rather than

[31] Although this use of *letterlijk* is listed as a separate sense in Van Dale, it does seem to be a source of some annoyance, with language enthusiasts calling it "erroneous", "irritating" or even "dangerous" (see, e.g., http://www.elsbethschrijft.nl/taal/letterlijk/, https://taaldacht.nl/2013/08/09/letterlijk/ https://irritaal.wordpress.com/2010/01/08/letterlijk/). As was already pointed out by Bolinger in 1972: 107-108, the English word *literally* is used in much the same way, as in e.g. *It literally took me ages to finish this assignment* or *I was literally dying with laughter*, and has incited similar reactions in the (online) media (see, e.g., https://www.theguardian.com/politics/reality-check-with-polly-curtis/2012/mar/12/reality-check-literally-wrong-use-word, http://www.npr.org/2014/05/25/315703164/a-literal-truce-over-the-misuse-of-literally, http://thewritepractice.com/stop-saying-literally/). Web pages last accessed on August 10, 2017.

treating the examples like (88) and (89) as truly ambiguous, i.e. as allowing for two different interpretations, we could say that they are vague. Semantic vagueness is theoretically different from ambiguity and polysemy in that vague meanings are unspecific or indeterminate and therefore difficult to separate (Zwicky & Sadock 1975). However, the model of three well-defined and clearly delineated categories (ambiguity, polysemy and vagueness) has been called into question because there are many examples of contexts in which the same set of meanings could be either separable or united, or both (Tuggy 1993, Geeraerts 1993). The distinction remains valid but we need to conceive of the categories as gradient rather than absolute. Applied to the fake reflexive resultative construction, we propose that there is a continuum between clear instances of resultative fake reflexives on one end (e.g. *Hij schiet zich dood* 'he shoots himself dead') and intensifying fake reflexives on the other (e.g. *Hij lacht zich rot* 'he laughs himself rotten'). The vague instances like (88) and (89) are situated somewhere in between the two poles and were not annotated as either intensifying or resultative. The number of such truly vague instances was fairly small: even in potentially vague (and/or ambiguous) cases, there are often still certain contextual clues that indicate whether a specific clause is *inching more towards* the intensifying or the resultative end of the continuum, see (90) to (92).

(90) Toen zij tijdens een bezoek aan een trainingscentrum voor huisbedienden zag hoe die **zich in het zweet werkten**, schonk zij het centrum een elektrische ventilator. (Delphcorp, 1980-1989)
[...] how these themselves in the sweat worked [...]
'When, during a visit to the training centre for house helps, she noticed how they were sweating while doing their work, she gave the centre an electrical fan.'

(91) Wanneer obers **zich in het zweet rennen**, de transpiratiegeur in de keukens de luchtjes van het vlees en de sauzen probeert te verdringen... (Delphcorp, 1980-1989)
[...] when butlers themselves in the sweat run [...]
'When butlers run around sweating and the smell of sweat in the kitchens tries to supplant the smell of the meat and sauces...'

(92) Daar **werkten** de bewindslieden **zich in het zweet** om moeizame compromissen tot stand te brengen. (Delphcorp, 1990-1995)
there worked the cabinet members themselves in the sweat [...]
'In there, the cabinet members worked very hard to reach tough compromises.'

In the examples above, we opted for a resultative reading in (90) and (91) because the mention of a cooling fan and the smell of sweat clearly puts the focus on the actual sweating. In (92) an intensifying reading seems more likely because the work that is mentioned here does not immediately involve heavy physical activity that may cause actual sweating, so the focus is most likely more on the hard work.

Because semantic annotation is often a matter of subtle distinctions and personal interpretation, we tested the validity of our categorisation of intensifying versus

resultative semantics with the calculation of inter-analyst agreement. Amélie Van Beveren, a colleague in the Dutch Linguistics department annotated 1,000 sentences (not including any of the truly vague or ambiguous instances, cf. (88) and (89) supra) that were randomly selected from both the synchronic and the diachronic data sets. We provided her with the above information on our classification, as well as some additional examples of both the resultative and intensifying categories and asked her to label each of the 1,000 sentences as either "INT" for intensifying or "RES" for resultative. She was made aware that the classification may not always be straightforward and that, in doubt, she had to choose the category which she found to be the "best fit". With κ=0.76, the agreement was substantial. After discussing the sentences that the annotators had originally labelled differently, we mutually came to an agreement on the annotation of these individual sentences. Figure 3.9 gives an overview of the frequency of the classifications in the SoNaR data set.

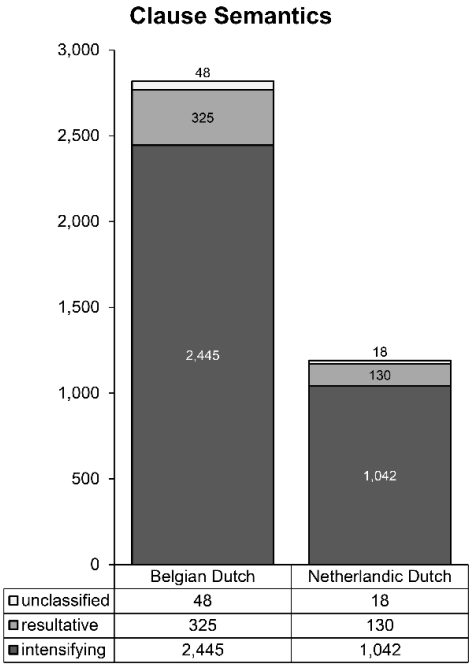


Figure 3.9. Proportion of the intensifying versus literal fake reflexive resultative construction (and unclassified vague/ambiguous cases) in SoNaR

All in all, the number of vague and/or ambiguous cases in which the linguistic context did not provide us with any immediate answers (like (88) and (89)) was fairly limited (not even 2% of the entire data set). The intensifying category outranks the resultative category by far – which is to be expected, as has been repeated throughout this chapter, given the bias in the search procedure. If we look at the frequency of the categories in Delphcorp in Figure 3.10 , we see very similar proportions in the late 20th Century, but the dominance of the intensifying category is much less pronounced in the 19th Century and the early 20th Century.

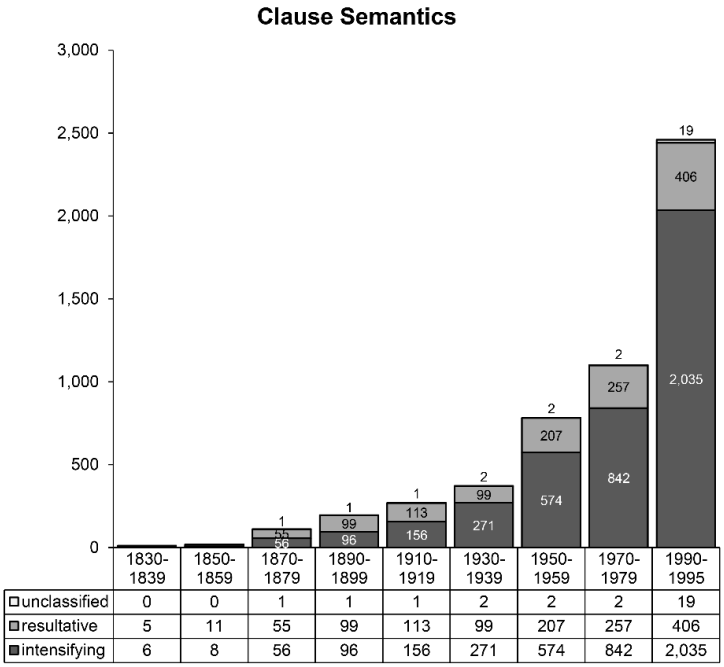


Figure 3.10. Proportion of the intensifying versus literal fake reflexive resultative construction (and unclassified vague/ambiguous cases) in Delphcorp

If the intensifying construction did arise out of the literal resultative construction, and if it was still in its infancy in the early 19th Century, we could expect the two constructions to behave more similarly than they do in present-day Dutch. In the next section, we will investigate whether this was indeed the case.

3.3.5.2 Diachronic development: divergence?

In the entire Delphcorp data set, we find 4,044 instances of the intensifying construction and 1,252 instances of the resultative construction. While these frequencies suggest a large dominance of the intensifying construction, the numbers are actually somewhat misleading in that the resultative instances only contain resultative phrases which were also found to function as intensifiers (cf. supra) – and therefore do not really represent the full extent of the literal fake reflexive resultative construction –, whereas the intensifying instances comprise all intensifiers, including the non-overlapping (i.e. exclusively intensifying) ones. It is interesting to note that we already find such non-overlapping intensifiers quite early on in the data, e.g. *een bult* 'a hump' in the 1850s or *het vuur uit de sloffen* 'the fire out of the slippers' in the 1870s.

(93) Och hou toch op!! Want ik **lach mij een bult**. (Delphcorp, 1850-1859)
[...] I laugh myself a hump
'Oh, stop it! I'm laughing so hard.'

Examples like (93) add further strength to the claim that the intensifying construction already existed as an independent construction at the time, as we would not expect to

find such instances if the intensifying interpretation was still pragmatically derived from the resultative construction at this point. If we want to investigate the potential ambiguity of the pattern, we can only take into account those lexical elements that were actually found to alternate between the resultative and the intensifying constructions, i.e. those elements that were attested as both a resultative phrase and an intensifier in the entire data set. Figure 3.11 gives the development of the relative proportions of both constructions limited to the subset of overlapping intensifiers (37 intensifiers in Delphcorp, 14 in SoNaR). In all bars, the first number is the normalised frequency per ten million words[32], the second the absolute frequency of occurrence.

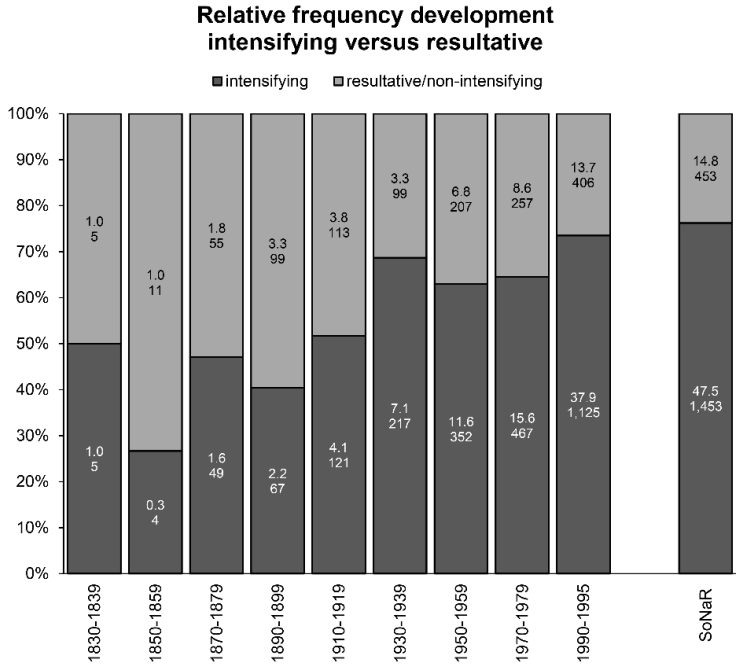


Figure 3.11. Relative frequency development of intensifying versus literal fake reflexive resultative construction for overlapping items

Although both constructions have expanded their use over time, the intensifying construction has done so at a much more accelerated rate: if we compare the frequencies from the late 19th Century in Delphcorp to the present-day Dutch data from SoNaR, the intensifying construction has become approximately 20 times more frequent whereas the resultative construction has only increased by a factor of 4.5.

Evidently, not all of the individual overlapping intensifiers show the same alternating behaviour or the same frequency development. For one, there are a number of items that are infrequent in both constructions – that is, they are neither strong intensifiers nor established resultative phrases –, the alternation of which therefore not being very

[32] We take ten million words as the normalising standard instead of the usual one million words, because the construction is overall rather infrequent, especially in the diachronic data.

informative. See, e.g., a resultative and intensifying example of *bewusteloos* 'unconscious' in (94) and (95), respectively.

(94) Een knaap van 17 **dronk zich** zaterdag **bewusteloos**. (Delphcorp, 1950-1959)
a lad of 17 drank himself saturday unconscious
'A lad of 17 years old passed out from drinking last Saturday.'

(95) Dan tik ik het onderwerp in op het scherm en dan krijg ik meteen het nummer van de band te zien. Anders **zoek** je **je bewusteloos**. (Delphcorp, 1990-1995)
[...] otherwise search you yourself unconscious
'If I type in the subject on the screen, I immediately get to see the number of the tape. Otherwise, you would have to search for hours.'

If we only consider the items that do appear to have some frequency in both constructions, we find that there are several cases in which the resultative use appears to predate the intensifying use in our data set, which is exactly what we would expect if the intensifying construction did arise out of the resultative construction. For example, even though *dood* 'dead' was already present as an intensifier in the construction in the 1830s, the fact that the resultative use was still dominant at the time indicates that the intensifying use of *dood* 'dead' is of a more recent date than its resultative use. A much more recent intensifier is *te pletter* 'to smithereens'. While *te pletter* was already used as a resultative phrase in the mid-19th Century, it was only introduced as an intensifier in the 1950s and did not really take off until the 1970s. If we look at the early intensifying examples of some of these intensifiers, we find that they are still somewhat reminiscent of the resultative use in that the verbs they co-occur with are also compatible with a (non-instantaneous) resultative reading.

(96) Zij zullen, zonder **zich** daarom **dood** te **werken**, de filtreer-dienst der drie directeuren, [...] kunnen waarnemen. (Delphcorp, 1830-1839)
they shall, without themselves therefore dead to work [...]
'They will be able to look after the filter-service of the three principals, without having to work too hard.'

It is quite likely that these intensifying uses have developed out of non-instantaneous resultatives through a shift in focus. In older literal resultatives with *werken* 'to work', for example, there may have already been an underlying degree aspect, in as far as the result can be construed as the outcome of the *intensive or excessive performance* of the verbal activity of working. It is possible that the focus on the end-result has weakened, thus paving the way for a degree meaning to arise as the central meaning of a new construction. The true experience verbs like *lachen* 'to laugh', *zich ergeren* 'to be annoyed',

etc. – which are much less readily conceivable in a resultative use[33] – generally appear later in this construction.

In present-day Dutch, there is often – though certainly not always – a clear difference with respect to the type of verbal collocates that are used with the intensifiers or with the resultative phrases. For example, whereas the intensifier *te pletter* 'to smithereens' occurs with a wide range of different verb types, the majority of the resultative tokens are accounted for by the collocations *zich te pletter lopen* 'to run into', *zich te pletter rijden* 'to crash into something when driving' (or occasionally 'to exhaust oneself by cycling/running', see the category of hyperbolic resultatives, supra) and *zich te pletter vliegen* 'to crash into something when flying'.

(97) Een massa van 150.000 Duitschers, uit nieuwe troepen gevormd [...] gaat **zich te pletter loopen** op het Belgisch leger. (Delphcorp, 1910-1919)
[...] going to itself to smithereens run on the belgian army
'A mass of 150,000 Germans, formed out of new troops [...] will run into (or be halted by) the Belgian army.'

While the resultative use may have been *temporally* primary, the intensifying use has developed into the "default" meaning in present-day Dutch in most combinations, in the sense that this is the interpretation that language users will immediately arrive at, as long as there are no contextual clues explicitly cancelling it (cf. supra). There are just a couple of items that still have a dominant resultative use even in present-day Dutch, such as *in het zweet* 'in the sweat'. There is a strong collocational overlap between the resultative and intensifying uses of *in het zweet*, which could be related to its rather specific semantics. Even in intensifying uses, we find activity verbs – mainly *werken* 'to work' – that are quite likely to cause sweating when performed with a certain intensity. In order to classify these examples, we heavily relied on textual context, as was explained in the previous paragraph. This could indicate that for this particular intensifier (and some others like it, e.g. *schor* 'hoarse'), the intensifying use is still heavily dependent upon the resultative use or that, at the very least, the resultative meaning still lingers in the background.

To be sure, even if the intensifying fake reflexive resultative construction in general may have originally developed from the literal fake reflexive resultative construction, it is not the case that the resultative use necessarily came first for *all* overlapping items. Language users may generalise the existing ambiguity of the [SUBJ V REFL XP] pattern (in sentences with, e.g., *dood* 'dead') to any item in the XP-slot, even to those that were introduced as intensifiers first – as long as a resultative interpretation is not de facto

[33] This is of course a matter of world knowledge. In theory, many activities could potentially cause death, exhaustion, drowsiness, etc., but the likeliness of this happening is much greater in the case of *werken* 'to work' than in the case of *zich ergeren* 'to be annoyed'.

impossible. This is what Margerie (2013) found to be the case for the English micro-constructions [NP BE SCARED *sick*] and [NP_1 SCARE NP_2 *sick*] (see also §2.2.1). Even though she concurs that the general degree modifier construction in English originally was reanalysed from a prior resultative construction – as is evident from her discussion of several patterns with *to death* in Margerie (2011) – she argues that for the specific low-level patterns with *sick*, the degree meaning actually originated prior to the resultative meaning. On the basis of early 20th-Century examples in the Old English Dictionary [OED], Margerie argues that the degree patterns [NP VB NP *rigid/stiff/silly*] or [NP BE NP *rigid/stiff/silly*] may have provided a model for analogical extension to *sick*, which has similar semantics. Such analogical developments, based on the similarity between these patterns, may have further entrenched the abstract subschemas. Once the degree modifier construction had become established, new resultative meanings for specific degree modifiers could arise as a result of the already existing ambiguity (viz. modelled on similar patterns the degree meaning of which *did* arise out of a resultative meaning, such as the pattern with *to death*). One of her arguments to support this hypothesis is the high frequency and cognitive salience of the degree meaning of *sick* in synchronic data. However, that in itself is not a sufficient argument because the degree/intensifying construction can of course develop into the most frequent or salient one even for those intensifiers that were originally resultative phrases (see, e.g., *dood* 'dead' and *te pletter* 'to smithereens' in Dutch). Coincidentally or not, the Dutch counterpart *ziek* 'sick' also appears to have been introduced as an intensifier *before* it was used as a resultative phrase. The intensifier *ziek*, in combination with the verb *lachen* 'to laugh', was already attested in the early 19th Century (and even before the 19th Century cf. the example from the WNT in Ch2, §2.2.2.3), see (98), but there are no clear resultative uses in the data until the 1870s, see (99).

(98) Wat zegt ge? – roept Spruit – Och, ik **lach me** nog **ziek**! – Ik zeg: met een stoomboot van gom-elastiek! (Delphcorp, 1830-1839)
[...] oh I laugh myself still sick [...]
'What did you say? Spruit yells. Oh I'm dying with laughter! – I said: with a steamboat made of gummi!'

(99) Men veronderstelt, dat de beesten **zich ziek** hebben **gevreten** aan den overvloed van 't jonge gras. (Delphcorp, 1870-1879)
[...] the animals themselves sick have stuffed [...]
'One supposes that the animals ate so much grass that it made them sick.'

Of course, it is hard to imagine that a sentence like (99) was "ungrammatical" or uninterpretable before the 1870s, but for whatever reason, there were no earlier resultative attestations with *vreten* 'to gorge oneself' or other verbs that are likely to cause sickness in our data set. There is no way to exclude the possibility that this is no more than a quirk in the data, but there are other examples which indicate that the resultative use may be derived from or secondary to the intensifying use. Take the "resultative"

example of *een ongeluk* 'an accident' in (100): it is not used in a full sentence but as a "catchy" headline, in which the journalist probably intended to play on the expressions *per ongeluk* 'by accident' and *zich een ongeluk V'en* 'to V oneself an accident'.

(100) **Zich een ongeluk slikken**: In de VS gaat het aantal kleine kinderen dat per ongeluk medicijnen inneemt en daardoor in een ziekenhuis belandt in stijgende lijn. (SoNaR-BE)
swallow oneself an accident [...]
'Swallowing (by accident) can cause accidents: In the US, the number of children that take medicine by accident and end up in a hospital is steadily increasing.'

The shared syntax of the intensifying and the literal resultative construction is not only an interesting object of study for the linguist; there are some indications that regular language users are aware of this formal surface similarity and the fact that it may give rise to ambiguity. As we showed in the previous paragraph, language users sometimes use explicit textual clues to guide the hearer towards the intended interpretation. Another possible indication that language users are aware of the ambiguity of the pattern is when they exploit it in order to ingeniously play with both meanings in the same sentence (see also Tuggy 1993 on puns involving the superimposition of two meanings). Example (101) invokes the scenario of some kind of strategic move in which the enemy is confronted with something extremely funny in order to make him laugh a lot (i.e. the intensifying interpretation of *hij lacht zich dood* 'he laughs himself dead'). Eventually, the enemy laughs so much that he literally dies (the resultative interpretation of *hij lacht zich dood* 'he laughs himself dead') and the war is won.

(101) Komt me daar de grens over om zo'n ellendig zoodje [sic] schandaal-lectuur binnen te smokkelen. Willen jullie dat iedereen aan deze kant van de grenslijn **zich dood lacht**? Is dat de opzet? (Delphcorp, 1950-1959)
[...] that everyone on this side of the border himself dead laughs? [...]
'So someone crosses the border to smuggle in a miserable pile of scandal sheets. Do you want everyone on this side of the border to die with laughter? Is that the intention?'

Summing up, although it is not our aim to delve into the ways in which the intensifying construction originally arose out of the resultative construction, the data do provide some evidence of increasing divergence between both constructions. This diverging movement could be interpreted as indirect evidence for a resultative origin, in that the intensifying construction appears to be gradually emancipating itself from the resultative construction. In present-day Dutch, the actual lexical overlap between resultative and intensifying constructions is fairly limited. The intensifiers that are exclusively used in the intensifying construction outnumber the alternating ones by far, and the intensifiers that do occur in both constructions generally behave differently in terms of the verbs they collocate with in both patterns.

In order to avoid overestimating the frequencies that will be used for the quantitative analyses in Chapters 4 and 5, we opted for the conservative route of only working with

the (more or less) clear intensifying examples in the remainder of this thesis. That is, we have excluded all unclassified items and all instances in which the postverbal phrase can still be seen as encoding some kind of (hyperbolic) end state, and we will only focus on the intensifying part of the synchronic and diachronic data sets.

3.3.6 Non-selected parameter: semantics of intensification

In Chapter 2, we briefly noted that the precise nature of intensification may show some variation depending on the element that is boosted and the context in which it is used. To round off this methodological chapter, we briefly want to illustrate what kind of variation is attested in the intensifying fake reflexive resultative construction. However, given that the distinctions are occasionally difficult to operationalise and require a certain degree of personal interpretation – as will become clear from the examples presented below – we have not included this semantic dimension as a real variable in the annotation.

According to the verbal dimension that is boosted, we could distinguish the following broad categories.

I. Intensity of experience: the subject has a heightened emotional or cognitive experience ⇔ to V intensely (to be very X-ed)

(102) De uitzending "Even afrekenen, heren", is meestal om **je groen en geel** aan **te ergeren**. (Delphcorp, 1950-1959)
[...] to yourself green and yellow at to annoy
'The show "Even afrekenen, heren" tends to make you very annoyed.'

(103) Ik durfde geen vriendinnetje mee naar huis te nemen omdat ik **me rot schaamde** voor mijn ouders. (SoNaR-NL)
[...] I myself rotten embarrassed for my parents
'I never dared bring one of my girlfriends home because I was so ashamed of my parents.'

(104) Ik **schrok me** dan ook **te pletter** toen Regine Beer me persoonlijk opbelde. (SoNaR-BE)
I startled myself then also to smithereens [...]
'I was naturally very startled when Regine Beer personally called me.'

II. Intensity of performance: the subject performs a (physical) activity with intensity ⇔ to V intensely, with a lot of effort

(105) De ploeg – niemand uitgezonderd – **werkte zich het snot voor de ogen**.(SoNaR-NL)
[...] worked itself the snot in front of the eyes
'The entire team, without exception, worked hard.'

(106) De deelnemers **trappen zich de longen uit het lijf**, maar komen nauwelijks vooruit. (Delphcorp, 1990-1995)
the participants pedalled themselves the lungs out of the body [...]
'The participants were pedalling like crazy, but they barely moved an inch.'

(107) Ben je gek, op deze muziek blijf je niet zitten, je **danst je uit de naad**. (Delphcorp, 1990-1995)
[...] you dance yourself out of the seam
'Are you crazy, it's impossible to sit still to this music, you're dancing your butt off.'

III. Duration/repetition: the subject performs the activity continuously or repeatedly over an extended period of time ⇔ to V extensively/often/for a long time

(108) En wanneer Dehaene, Kohl en Prodi **zich het pleuris vergaderden** om hun land de normen van Maastricht te doen slikken, ontdekte senator Verhofstadt de wondere werelden van Mario Vargas Llosa en James Joyce. (SoNaR-BE)
and when dehaene, kohl and prodi themselves the pleurisy meet [...]
'And while Dehaene, Kohl and Prodi hold meeting after meeting to shove the norms of Maastricht down their countries' throats, senator Verhofstadt discovered the wondrous worlds of Mario Vargas Llosa and James Joyce.'

(109) Binnen de kortste keren doorwoelt een kolonne van cliënten het echtelijke bed. "Ik **neuk me lam**", zegt Pipo. (Delphcorp, 1990-1995)
[...] I fuck myself lame, says pipo
'In no time, a bunch of clients rummage through the conjugal bed. I fuck a lot, Pipo says.'

(110) Leuk, maar je hebt er Flash voor nodig, alles laadt erg langzaam en je zit **je suf** te **klikken** om bij de echte informatie te komen. (SoNaR-NL)
[...] you sit yourself drowsy to click [...]
'It's fun, but you need Flash to run it; everything takes a long time to load and you have to keep clicking to get to the real information.'

IV. Direct object: the intensifier does not really boost an inherent property of the verb, but relates to the direct object ⇔ to V a large amount of X

(111) Of in de wijk Shi Men Ding, waar de trendgevoelige Taiwanees **zich ongans koopt** aan alles wat een Japanse techno-uitstraling heeft. (SoNaR-NL)
[...] the fashionable taiwanese himself unwell buys [...]
'Or in the district Shi Men Ding, where fashionable Taiwanese buy everything that looks like Japanese technology.'

(112) Hij dacht waarschijnlijk aan zijn arme landgenooten, die **zich** thans **blauw betalen** aan belastingen. (Delphcorp, 1930-1939)
[...] who themselves now pay blue [...]
'He was probably thinking about his poor compatriots, who are now paying a lot of taxes.'

(113) Wie dan door wil gaan voor een Mr- of Drs- titel zal **zich gek** moeten gaan **lenen** bij de bank. (SoNaR-NL)
[...] will himself crazy have to go loan [...]
'Whoever wants to go for a Mr- or Drs-title will have to loan a lot of money from the bank.'

V. Verb-specific dimensions: the intensifier boosts a dimension that is specific to a particular verb type ⇔ to V loudly/fast/...

(114) Waarna ze vijf minuten later weer aanhaakten bij "Whole lotta love", en weer **juichte** het publiek **zich de ziel uit het lijf**. (SoNaR-BE)
[...] cheered the audience itself the soul out of the body
'After five minutes they picked up again at "Whole lotta love", and another loud cheer went through the crowd.'

Fundamentally, the categories do not a priori delineate clusters of individual verbs or intensifiers and it is important to apply the categorisation on an item-by item basis. For example, the verb *rennen* 'to run', can occur in multiple categories, as demonstrated by the following examples. In (115), it is the running-inherent dimension of speed that is highlighted, whereas (116) expresses the notion of someone constantly running around, without there necessarily being much speed involved.

(115) Binnen vijf minuten werden er een kilometer verderop in de Kinkerstraat ook ruiten ingegooid. Je moet **je** echt **te pletter rennen**, om dat te kunnen voorkomen. (SoNaR-NL)
[...] you have to yourself really to smithereens run [...]
'Five minutes later, windows are also being smashed about a kilometre up ahead in the Kinkerstraat. You would have to run really fast to prevent that.'

(116) Geüniformeerde obers **rennen zich** er **rot** met koffie en gebak voor een zeer gemengde clientele. (SoNaR-NL)
[...] run themselves there rotten [...]
'Waiters in uniforms run around like crazy with coffee and cake for a very diverse group of customers.'

Another example is the intensifier *blauw* 'blue'. In combination with the verb *betalen* 'to pay', it adds the meaning of 'to pay a lot', see (112) (category IV), but in combination with the verb *zich ergeren* 'to be annoyed', it expresses that the subject is experiencing heightened annoyance (category I). In example (117), *blauw* highlights the frequency with which the media are writing protest letters or possibly the amount of protest letters that have been written (categories III and IV).

(117) Velerlei mediaorganisaties hebben **zich** de afgelopen jaren **blauw geschreven** aan protestbrieven. (SoNaR-NL)
[...] have themselves blue written [...]
'A lot of media companies have been (frequently) writing (a lot of) protest letters over the past couple of years.'

Overall, the majority of the data are rather clear instances of either intensity of experience or intensity of performance, but some of the examples above show that the categories are by no means mutually exclusive and may even coincide in one and the same sentence. Indeed, if someone is running very fast, they are at the same time running with a certain intensity (see (115), categories II and V) or if someone often writes letters, this necessarily implies that they write a lot of letters (see (117), categories III and IV). Given the interrelatedness of some of the categories, we will take all of the dimensions together under the general notion of intensification for present purposes.

Chapter 4
Synchronic use and variation: the intensifying fake reflexive resultative construction in present-day Netherlandic and Belgian Dutch

In this chapter, we will compare the use of the intensifying fake reflexive resultative construction in Netherlandic and Belgian Dutch by looking into (i) the general use and frequency of the construction, (ii) the collocational patterns, (iii) degrees of productivity and (iv) the hierarchic organisation of the constructional network. Each aspect is discussed in a separate section, which is in turn divided into two parts. In the first part, we present a detailed picture of the synchronic use in Netherlandic Dutch; in the second part of each section, we look at the synchronic variation by comparing Netherlandic Dutch and Belgian Dutch. The first part of the synchronic use will then serve as a point of comparison for the investigation of diachronic variation that is presented in the next chapter. In §4.1, the aim is to sketch a general picture of the use of the intensifying fake reflexive resultative construction in present-day Dutch. Concretely, we will look at the overall frequency of the construction and analyse its different components to document how speakers of Dutch fill in the empty slots of the construction. Going beyond the individual slots in §4.2, we consider the possibility of coselection or covariation, in which the filler of one slot (partly) determines the filler of another slot, in order to examine whether there are any notable consistencies in the way in which specific verbs and intensifiers are linked. In particular, we are interested to learn whether it is possible to distinguish conventionalised verb-intensifier collocations from on-the-fly, creative verb-intensifier combinations. As will be shown, a close inspection of the collocational patterns naturally brings us to the topic of productivity, which is discussed in section §4.3. This section aims to measure constructional productivity by applying a multidimensional model of productivity (cf. Chapter 2). Particular attention is paid to the question as to how productivity should be interpreted at different hierarchical levels and how it shapes the

internal structure of the constructional network, which will be further discussed in §4.4. Finally, the main findings of this chapter will be summarised in §4.5.

4.1 A preliminary look at frequency data and slot fillers

4.1.1 Synchronic use

The Netherlandic part of the SoNaR data set [=SoNaR-NL] contains 1,042 unambiguous instances of the intensifying fake reflexive resultative construction. Taking into account the corpus size (cf. Chapter 3), we get a normalised frequency of approximately 142.5 instances per ten million words. In what follows, we will first have a more thorough look at the way in which speakers of Netherlandic Dutch fill in the specific slots of the pattern [SUBJ V REFL INT]. This should already give us a first indication of the overall usage pattern of the intensifying fake reflexive resultative construction in present-day Dutch. Figure 4.1 summarises the frequency information for the variables Reflexivity, Transitivity, Syntactic Category and Reflexive Pronoun that were discussed in Chapter 3. In the following paragraphs, we will discuss each of the panels in more detail.

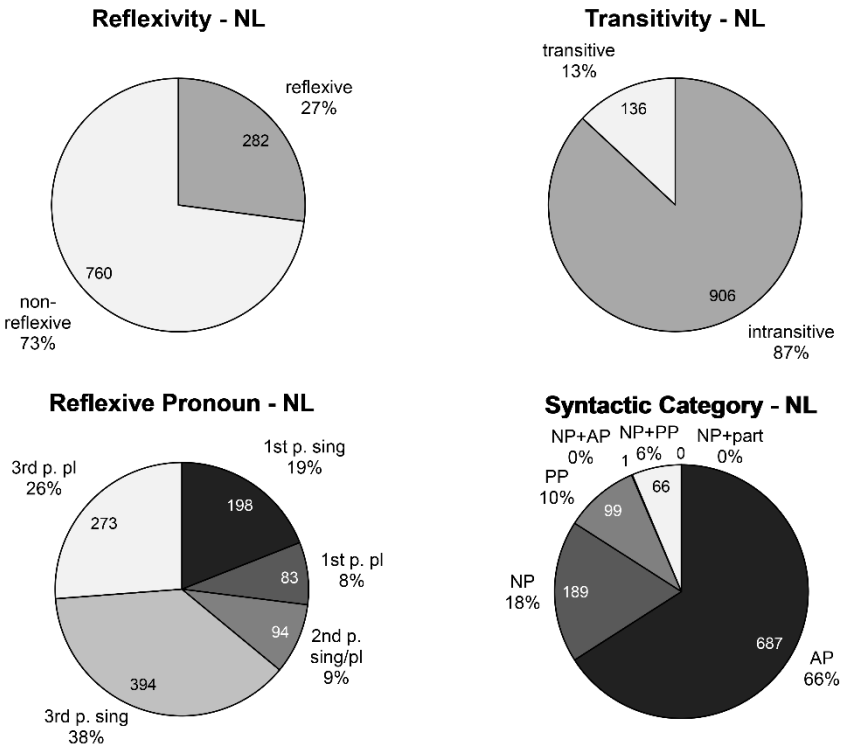


Figure 4.1. Summary of the variables in SoNaR-NL

4.1.1.1 Verb

In all, the Netherlandic data contain 137 different verb types. The top ten of the most frequently used verbs in this construction in Netherlandic Dutch are listed in (v), with their token frequencies between brackets (see Appendix IV-1 for the full list of verbs).

(v) Top ten verbs in SoNaR-NL

1. schrikken 'to be startled'	(223)
2. **zich ergeren 'to be annoyed'**	(133)
3. werken 'to work'	(112)
4. **zich schamen 'to be embarrassed'**	(80)
5. lachen 'to laugh'	(70)
6. **zich vervelen 'to be bored'**	(58)
7. lopen 'to run'	(36)
8. piekeren 'to worry'	(22)
9. betalen 'to pay'	(18)
10. zoeken 'to search'	(14)

The overview of top verbs shows that although the 1,042 tokens are distributed over 137 verb types, a large part of the set is accounted for by a small number of very frequent verbs. This Zipfian-like distribution is characteristic of the token frequencies of words in constructional slots (Ellis & Ferreira-Junior 2009, Gries 2012). If we combine the upper panels of Figure 4.1, and take into account that inherently reflexive verbs are necessarily intransitive, we find that the lion's share of the items in the data set (i.e. 624 in total) features non-reflexive, intransitive verbs. Nevertheless, the inherently reflexive verbs play a more crucial role in the construction than the circle diagram gives them credit for. This becomes clear if we do not only consider the token frequency, but also the type frequency of the different verbal categories.

Table 4.1. Verb proportion (reflexivity and transitivity combined) in the intensifying set in SoNaR-NL

	TOKENS	TYPES
Transitive verbs	136	39
Intransitive non-reflexive verbs	624	90
Intransitive reflexive verbs	282	8
TOTAL	1,042	137

Table 4.1 shows that the 282 reflexive tokens feature only 8 inherently reflexive verbs. More precisely, looking at the actual data, we see that 96% of the reflexive tokens, or 26% of the entire data set, are accounted for by just 3 inherently reflexive verbs, which also feature among the top ten verbs overall (see the bolded verbs in (v)). Earlier, we assumed that there might be a difference between reflexive and non-reflexive verbs, based on the

fact that the inherently reflexive verbs already subcategorise for a reflexive pronoun, which merges with the constructional reflexive pronoun, whereas the non-reflexive verbs do not (see Ch3, §3.3.2). Later in this section, we will investigate whether there are any behavioural differences between reflexive and non-reflexive verbs with respect to their collocational patterns and combinatorial flexibility. Overall, most of the verbs that occur with a certain frequency in the construction seem to belong to a limited number of broadly defined semantic clusters. First of all, there is a large category of "experience verbs", which can be split up into two subcategories: (A) verbs of emotion and (B) mental or cognitive (activity) verbs.

I. A. Experience (emotion verbs)*: lachen* 'to laugh', *schrikken* 'to be startled', *zich amuseren* 'to enjoy oneself', *zich ergeren* 'to be annoyed', *zich schamen* 'to be embarrassed, *zich vervelen* 'to be bored'...

B. Experience (cognitive verbs): *denken* 'to think', *lezen* 'to read', *peinzen* 'to ponder', *piekeren* 'to worry'...

The frequency of occurrence of the inherently reflexive verbs could be related to their semantics, rather than their inherent reflexivity: just like the frequent non-reflexive verbs *lachen* 'to laugh' and *schrikken* 'to be startled', the reflexive verbs express a basic human experience that is inherently prone to intensification (cf. Ch2, §2.3). While these are among the most token frequent verbs overall, the category of experience verbs is not extremely type frequent. The second category is more varied in terms of different types, while also containing a number of highly token frequent verbs. This category consists of the verbs that require some physical effort, and which can therefore be performed with a certain intensity.

II. Physical activity: *dansen* 'to dance', *fietsen* 'to cycle', *lopen* 'to run', *sjouwen* 'to drag', *trainen* 'to train', *trappen* 'to pedal', *werken* 'to work', *zwemmen* 'to swim'...

We also find multiple verbs of communication in the construction, including verbs denoting new means or instruments of communication like emailing or texting.

III. Communication verbs: *bellen* 'to call', *discussiëren* 'to discuss', *kletsen* 'to chatter', *mailen* 'to email', *onderhandelen* 'to negotiate', *praten* 'to talk', *schrijven* 'to write', *sms'en* 'to text'...

Related to the communication verbs are the verbs of sound emission, a small and overall infrequent group.

IV. Sound emission: *blazen* 'to blow/play an instrument', *gillen* 'to screech', *schreeuwen* 'to scream', *spelen* 'to play an instrument', *zingen* 'to sing'...

Another small, yet coherent, semantic category are the verbs of consumption.

V. Consumption: *consumeren* 'to consume', *drinken* 'to drink', *eten* 'to eat', *roken* 'to smoke', *snuiven* 'to sniff drugs', *zuipen* 'to booze'...

We also encounter a large variety of different types of verbs that resist straightforward categorisation in one of the defined semantic classes, especially among the lower frequency verbs. This remaining category mainly contains activities that generally do not require extreme effort (unlike the activities in category II).

VI. Others/general activity: *aaien* 'to stroke', *bezuinigen* 'to economise', *lijnen* 'to diet', *lobbyen* 'to lobby', *printen* 'to print', *registreren* 'to register', *reizen* 'to travel', *turven* 'to tally'...

Clearly, a large variety of verbs can be boosted when used in this construction, although the exact nature of intensification can be somewhat different depending on the verb and the specific context in which the verb is used (cf. Ch3, §3.3.6). Still, we would not want to claim that the verb slot is entirely schematic in the sense that any verb could be used in the construction. The verb slot is semantically constrained insofar as the verb has to express some kind of experience or activity which has an inherent property that can be intensified in one way or another (cf. the aspects of intensification in Ch3, §3.3.6). Stative or durative verbs like *wonen* 'to live', *bestaan* 'to exist', *liggen* 'to lie', *staan* 'to stand', etc. are much less compatible with an intensifying reading so these do not readily fit into the construction. Also unlikely to occur in the construction are typical unaccusative verbs, in which the subject does not control the action of the verb, e.g. *bevriezen* 'to freeze', *breken* 'to break', *smelten* 'to melt', *sterven* 'to die', *stinken* 'to smell bad', *vallen* 'to fall', etc., although there do appear to be some exceptions here. Example (118) below, for instance, although admittedly rather poetic, contains an unaccusative verb with an inanimate subject, but we still understand that an intensifying meaning is conveyed.

(118) De zon zit klem in het woud. Ze **schittert zich te pletter** waar de takken wijken. (SoNaR-NL)
[...] she shines herself to smithereens [...]
'The sun is stuck in the woods. She shines intensely through the branches.'

In addition, there are various other verbs that sound odd in the construction (e.g. *leggen* 'to lay', *vinden* 'to find', *zetten* 'to put'...) because they are not easily conceived of as being performed with a certain intensity.

4.1.1.2 Reflexive pronoun

The lower left panel reveals that the intensifying fake reflexive resultative construction is most frequently used in the third person. This preference is robust across the different

categories of the data set (viz. reflexivity, transitivity, syntactic category). In light of the expressive and subjective nature of the construction, we might have expected to find it primarily used in the first person (singular), that is, when people are talking about their own experiences. However, we have just seen that the intensifying fake reflexive resultative construction is not just used to talk about personal "experiences" (though even in those cases the third person is more frequent), but that it may be used to boost a large variety of different activities. This may indicate that this construction is not just used as a way of exaggerating one's own experiences, but as a conventional expression of verbal intensification. There may be another reason why the third person clauses are abundant, related to the nature of the corpus we are working with: the main goal of journalism is to report about events that have happened to other people. Despite the majority of third person pronouns, we still find a considerable proportion of first and second pronouns as well. This may be related to the use of direct speech, which has become an important tool in journalism over the past decades (cf. Ch3, §3.1.1).

4.1.1.3 Intensifier

The Netherlandic Dutch data set features 68 different intensifier types, 23 of which are hapaxes. The top ten of the most frequently used intensifiers are given in (vi) below (see Appendix IV-2 for the full list of intensifiers).

(vi) Top ten intensifiers in SoNaR-NL

1. **rot 'rotten'**	(154)
2. **suf 'drowsy'**	(123)
3. **dood 'dead'**	(116)
4. **kapot 'broken'**	(104)
5. een ongeluk 'an accident'	(60)
6. te pletter 'to smithereens'	(60)
7. **groen en geel 'green and yellow'**	(44)
8. een slag in de rondte 'a punch around'	(37)
9. **wild 'wild'**	(34)
10. **blauw 'blue'**	(28)

In this section we will examine the formal properties of these intensifiers, i.e. the different syntactic categories they represent, and their semantic properties. In Chapter 2, §2.3, it was mentioned that cross-linguistically, intensifiers appear to be drawn from a limited number of conceptual source domains. We will see whether this also applies to the intensifiers that are used in this particular construction.

(a) Syntactic categories

Focusing our attention on the lower right panel of Figure 4.1, we see that the adjectival phrases largely outnumber all other categories in terms of token frequency. However,

this does not tell us how the 68 intensifier *types* are distributed across the syntactic categories, which is perhaps the more interesting question.

Table 4.2. Frequency comparison syntactic categories of the intensifiers in SoNaR-NL

	TOKENS	TYPES	INT. TTR	HAPAX TYPES	VERB RANGE
AP	687	25	0.036	6	97
NP	189	24	0.127	11	50
PP	99	4	0.040	0	28
NP+PP	66	14	0.212	5	25
NP+AP	1	1	1.000	1	1
TOTAL	1,042	68	0.064	23	137

Comparing the token and type counts in Table 4.2, we can add that although the adjectival phrases are best represented overall, they do not score high in terms of lexical variation, as measured by the type-token ratio. The AP category also has a relatively low number of hapaxes, compared to the NP and NP+PP categories. It would seem, then, that the individual adjectival intensifiers have a higher frequency of use overall, and that there are proportionally more high frequency adjectival intensifiers in the data set. For one, there are 9 adjectival intensifiers with 10 or more total occurrences, compared to only 3 prepositional intensifiers, 3 nominal intensifiers, and 2 NP+PP intensifiers.

(vii) AP: *ongans* 'unwell' (11), *wezenloos* 'blank/vacant' (22), *blauw* 'blue' (28), *wild* 'wild' (34), *groen en geel* 'green and yellow' (44), *kapot* 'broken' (104), *dood* 'dead' (116), *suf* 'drowsy' (123), *rot* 'rotten' (154)

(viii) PP: *in het zweet* 'in the sweat' (10), *uit de naad* 'out of the seam' (21), *te pletter* 'to smithereens' (60)

(ix) NP: *een hoedje* 'a little hat' (28), *een slag in de rondte*[34] 'a punch around' (37), *een ongeluk* 'an accident' (60)

(x) NP+PP: *de benen PREP het lijf* 'the legs PREP the body' (11), *de longen uit het lijf* 'the lungs out of the body' (12)

In addition, among the top ten intensifiers overall, there are 7 intensifiers from the adjectival category, indicated in bold in (vi). The top four intensifiers all have more than 100 individual tokens and together already cover approximately 71% of the tokens in the AP category or almost half (47%) of the entire SoNaR-NL data set.

Another way of comparing the syntactic categories is by looking at their verbal ranges (the rightmost column in Table 4.2). The adjectival intensifiers are found with 97 out of the 137 different verb types, which tells us that 70% of the verbs in our data set can occur

[34] *Een slag in de rondte* may superficially look like an NP+PP, but it is actually more like a complex NP. The prepositional part *in de rondte* is an attributive adjunct that says something about the head noun of the nominal phrase *slag*.

with at least one of the 25 adjectival intensifiers – note, though, that there is no one-to-one relationship and many verbs are found to combine with intensifiers from different categories. Curiously, if we tease the reflexive and non-reflexive verbs apart, we see that the overall dominance of the adjectival intensifiers is even more pronounced in the inherently reflexive category, see Figure 4.2. Especially the low number of NP and NP+PP intensifiers in combination with reflexive verbs is striking.

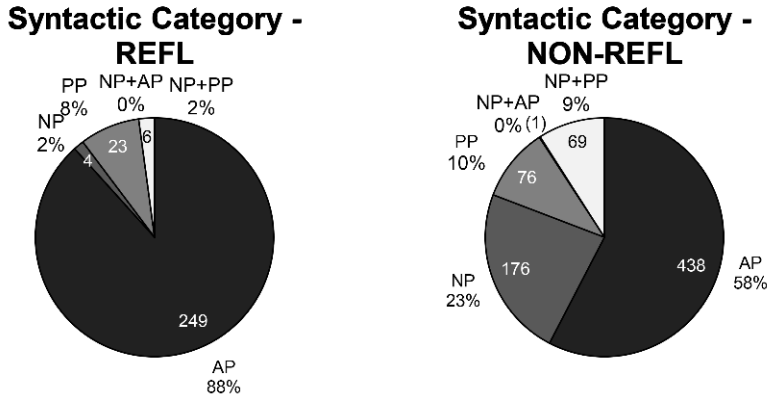


Figure 4.2. Comparison of the syntactic category of intensifiers combining with reflexive versus non-reflexive verbs in SoNaR-NL

The frequency of occurrence and combinatorial flexibility of the adjectival category raises the question as to what exactly sets this category apart from the other intensifier categories. In Chapter 2, §2.3, it was already pointed out that that the bulk of the work on degree modification deals with adjectival and adverbial modifiers, while studies focusing on other patterns of intensification are relatively rare. It may be the case that adjectives are more sensitive to bleaching or more inclined to develop intensifier functions than some of the other syntactic categories. As some of the adjectives are already polysemous to a certain degree in their lexical sense (e.g. *kapot* can mean 'dead', 'broken' or 'exhausted', which are listed as separate senses of the adjective in Van Dale), the step towards adopting another new meaning may be smaller than for very specific NPs like *een hoedje* 'a little hat' or "lexically heavy" NP+PP intensifiers like *de longen uit het lijf* 'the lungs out of the body'. Still, only a couple of the top adjectival intensifiers also have intensifying uses outside of the intensifying fake reflexive resultative construction (e.g. dood 'dead', kapot 'broken' and ziek 'sick' can be used as part of elative compounds like doodgemakkelijk 'lit. dead-easy', kapot mooi 'lit. broken pretty', ziek goed 'lit. sick good', see Hoeksema 2012, Ten Buuren et al. to appear), so it is definitely not the case that the adjectives are recognised as general, conventionalised intensifiers in Dutch. Moreover, we also find expressive uses for some of the nominal intensifiers in the construction: there is some overlap with the nominal elements that can occur in the *Krijg-de*-X-construction, e.g. *Krijg de klere/tyfus/vinkentering/pleuris/...* 'Go to hell/bugger off' (see also Ch2, §2.3 and the references therein). In Chapter 2, it was further mentioned that many of the lexical items that have developed intensifying properties are recruited from a

number of distinct source domains, which are characterised by shared inherent negativity. In the next paragraph, we will investigate whether we recognise some of these categories in the intensifiers in the intensifying fake reflexive resultative construction.

(b) Semantic categories

A closer look at the 68 intensifier types that occur in the construction reveals a number of semantic classes or conceptual domains from which multiple intensifiers in the intensifying fake reflexive resultative construction have been recruited. There is a large group of adjectival intensifiers which could be understood to denote an undesired emotional or cognitive state, *dood* 'dead' naturally being one of them.

I. Negatively connoted states*: bewusteloos* 'unconscious', *blind* 'blind', *dood* 'dead', *gek* 'crazy', *halfdood* 'half-dead', *kapot* 'broken', *klem* 'stuck', *krom* 'bent', *lam* 'lame', *ongans* 'unwell', *rot* 'rotten', *slap* 'weak', *wezenloos* 'blank/vacant', *wild* 'wild', *ziek* 'sick'...

The prepositional intensifiers *te pletter* 'to smithereens', *uit de naad* 'out of the seam' and *te barsten* 'to bursts' could also be added to the list in (I) as less prototypical members, to the extent that they also express some kind of unwanted state. A typical feature of Dutch that was mentioned in Chapter 2 is the intensifying use of diseases, in particular eradicated diseases that used to be very lethal or even fictitious diseases that are often modelled on existing diseases. Indeed, the list in (II) shows that these diseases form an important group in the intensifying fake reflexive resultative construction. Especially striking about this category is the *variety* of different diseases; most of these intensifiers are not very token frequent per se.

II. Diseases*: een delirium* 'a delirium', *een rolberoerte* 'a fit', *de kolere* 'the cholera', *de pest* 'the plague', *de pleuris* 'the pleurisy', *de tering* 'the consumption', *het apelazerus* 'fictitious disease', *het apenzweet* 'fictitious disease', *het apezuur* 'fictitious disease', *het lazerus* 'the leprosy', *het leplazerus* 'fictitious disease', *het schompes* 'fictitious disease'...

Although *een breuk* 'a fracture', *een bult* 'a hump' or *een kriek* (used in its archaic sense of a hump, not as a sour cherry) are not really diseases in the strict sense of the word, we could add them to the diseases group as they do denote some kind of physical ailment. In a way, this also applies to the intensifiers in (III) below, but we have put these in a separate group as they all explicitly involve an inalienable body part (or, occassionally, a piece of clothing).

III. Inalienable possession:[35] *de benen PREP het lijf* 'the legs PREP the body', *de benen PREP het gat* 'the legs PREP the butt', *de blaren op de tong* 'the blisters on the tongue', *de longen uit het lijf* 'the lungs out of the body', *de ogen uit het hoofd* 'the eyes out of the head', *het vuur uit de sloffen* 'the fire out of the slippers', *het vel van de botten* 'the skin off the bones', *de nieren los* 'the kidneys loose'...

As some of the intensifiers in this category express a relation of removal, they are reminiscent of the Body-Part-*Off* construction, see Chapter 2, §2.2.1. Occasionally, we also find NP+PP intensifiers in which the PP is *added* to the NP, e.g. *de blaren op de tong* 'the blisters on the tongue', as well as one NP+AP intensifier in which the body part is "affected" in a more general sense, viz. *de nieren los* 'the kidneys loose'. So far, the data confirm the overall propensity towards negatively evaluated items. A group that was not yet mentioned in Chapter 2, and which deserves some special attention here, are the colour terms in (IV).

IV. Colour terms: *blauw* 'blue', *bont en blauw* 'black and blue', *groen en geel* 'green and yellow', *paars* 'purple'...

There is something special about colour terms in that there is a well-documented cross-linguistic tendency to relate certain emotions to specific colours, although the exact associations can differ from one language to another (see Adams & Osgood 1973, Ogarkova 2007, Clarke & Costall 2008, Soriano & Valenzuela 2009, inter alia). For example, in German the colour of envy is yellow (*Gelb vor Neid* 'yellow with envy'), whereas in English and Dutch envy is associated with the colour green (*groen van jaloezie, green with envy*). In French, however, green is the colour of fear (*vert de peur* 'green with fear'), or can be associated with anger (*vert de rage* 'green with rage'), just like in Italian (*verde di colera* 'green with rage') (Soriano & Valenzuela 2009: 422). This inherent emotional value of colour terms may have contributed to their development as intensifiers. Two of the items in (IV) also have particular uses in other fixed expressions in Dutch: *bont en blauw* 'black and blue' is mainly used in the sense of beating someone black and blue, and *groen en geel* 'green and yellow' is also used in the expression *iemand groen en geel voor de ogen worden* 'his head began to swim'. Finally, we find a number of intensifiers which do not belong to any of the obvious conceptual domains. We list a couple of these "isolated" intensifiers in (V).

[35] The PREP-notation in some of these intensifiers indicates that different prepositions of removal were found in this position. In Chapter 3, §3.3.3, we explained why we have decided to conflate these different types into one lemma.

V. Others: *een eind in de rondte* 'a distance around', *een hoedje* 'a little hat', *een ongeluk* 'an accident', *in het zweet* 'in the sweat', *een slag in de rondte* 'a punch around', *een ootje* 'an old lady (?)', *een zoeavenmuts* 'a Zouave-bonnet'...

Interestingly, some of the hapax intensifiers belong to one of the established semantic domains, suggesting that the domains themselves are productive and sometimes recruit new intensifiers (cf. §4.4 infra). In (119) we see a hapax colour term and in (120) a hapax (fictitious) disease.

(119) De tweede vreest voor zijn privileges, maakt af en toe wat excuses, **manipuleert zich paars**. (SoNaR-NL)
[...] manipulates himself purple
'The second one is scared of losing his priviliges, comes up with excuses every now and then and manipulates the hell out of people.'

(120) Ik was gewend om **me het apenzweet** te **werken** in een onverwarmde kelder die we bij de gratie Gods van jezuïeten mochten gebruiken. (SoNaR-NL)
I was used to myself the monkey-sweat to work [...]
'I was used to working very hard in a cold basement that we were allowed to use, by the grace of God, by the Jesuits.'

Occasionally, we find a hapax intensifier which is modelled on a highly frequent existing intensifier that does not necessarily belong to one of the more type frequent semantic categories. In example (121) below *zich een zoeavenmuts schrikken* 'to startle oneself a Zouave-bonnet' is likely a variation on *zich een hoedje schrikken* 'to startle oneself a little hat'. It is possible that the conventional expression is slightly modified, by replacing *een hoedje* with a more specific type of headwear, in order to draw the reader's attention or to create some extra effect.

(121) Zo zijn Nederlanders: eerst rennen ze als kippen zonder kop, daarna **schrikken** ze **zich een zoeavenmuts**. (SoNaR-NL)
[...] startle they themselves a zouave-bonnet (i.e. a bonnet worn by the Zouaves)
'That's so typical of Dutchmen: first they run around like headless chickens and then something scares the hell out of them.'

Hapaxes may also be totally creative "semantically isolated" intensifiers, see (122), which suggests that the intensifier slot at the highest, most abstract level has some degree of "openness" and that speakers of Dutch are aware of the creative possibilities of this construction.

(122) Hij stelde voor het pasje het satirepasje te noemen. Ik heb **me een ootje gelachen**. (SoNaR-NL)[36]
[...] I have myself an 'ootje' laughed
'He suggested calling the permit the 'satire-permit'. I had a good laugh.'

The question now arises as to whether the original meaning of the lexical items may still be of importance when they are used in this particular construction. In other words, to what extent do the lexical semantics still persist (cf. the principle of persistence in grammaticalisation research, Hopper 1991) and in what way does this persistence manifest itself? One indication that the original lexical semantics are still present in the background is found in the phenomenon of one intensifier "outclassing" another intensifier. There are some indications that the strength of intensifiers is to some extent linked to their original semantics (in addition to their novelty; see Ch2, §2.3 on why novel intensifiers are felt to be more expressive or have a stronger booster effect). In (123), we get the impression that someone who 'startles himself rotten' is somehow less startled than someone who 'startles himself dead', which is in accordance with the fact that *dood* 'dead' is a more definite state than *rot* 'rotten'.

(123) Vervolgens **schrik** je **je rot** (toch **niet dood** hè). (SoNaR-NL)
then startle you yourself rotten (yet not dead eh)
'Then you are very scared (but not scared to death, right).'

Similarly, working *drie/vijf slagen in de rondte* 'three/five punches around' may be stronger than working *een slag in de rondte* 'one punch around' because three and five are mathematically more than one.

(124) Het college in Rotterdam **werkt zich drie slagen in de rondte** om uitvoering te geven aan de plannen. (SoNaR-NL)
the college in rotterdam works itself three punches around [...]
'The college in Rotterdam works very hard to put the plans into effect.'

In the section on productivity below, we will investigate how the semantic persistence effect may also explain certain distributional constraints.

[36] According to Van Dale, *een ootje* is a synonym for old lady but it may also refer to the circle that children sit in during several children's games. However, the term is better known for its use in the expression *iemand in het ootje nemen*, which means to trick or fool someone. It is likely that this expression has contributed to its being used as an intensifier in the intensifying fake reflexive resultative construction.

4.1.2 Synchronic variation

The Belgian Dutch data set consists of 2,445 intensifying clauses, which is more than double the size of the Netherlandic data set. However, if we factor in the big difference in corpus size, we find that the intensifying fake reflexive resultative construction is overall slightly less frequent in Belgian Dutch, with a normalised frequency of 105.2 instances per ten million words versus 142.5 instances per ten million words in Netherlandic Dutch (χ^2=67.89, p<.0001). In this section, we will examine whether the national varieties also show differences with respect to the concrete instantiation of the empty slots in the construction. The verb slot and the intensifier slot, as well as their interactions (§4.2), are especially instructive in uncovering whether speakers of Belgian Dutch and Netherlandic Dutch display different preferences. We first give a summary of the frequency information for the variables Reflexivity, Transitivity, Syntactic Category and Reflexive Pronoun in Figure 4.3. In the following paragraphs, we will compare the panels of Figure 4.3 and Figure 4.1 in detail.

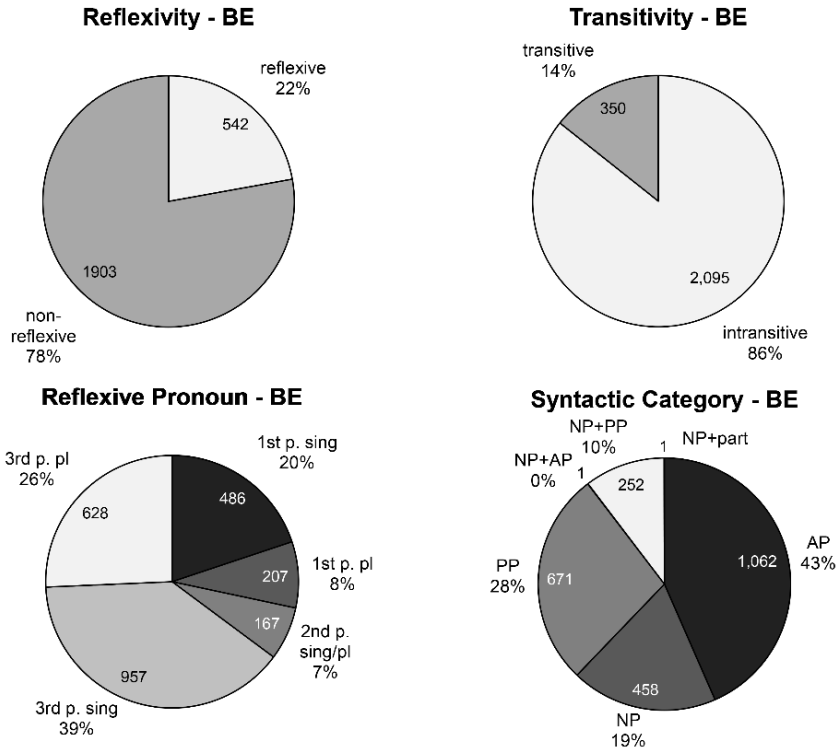


Figure 4.3. Summary of the variables for the intensifying set in SoNaR-BE

4.1.2.1 Verb and reflexive pronoun

With 167 verb types in 2,445 tokens (TTR=0.07), the Belgian Dutch data set shows a lower degree of variation of verbs overall than the Netherlandic Dutch data set (137 types in 1,042 tokens; TTR=0.13). A side-by-side comparison of the top ten verbs in (xi) shows an overlap of 8 verbs, indicated in bold font, suggesting that speakers of both national varieties tend to boost the same verbs.

(xi) Top 10 verbs in Belgian and Netherlandic Dutch

BELGIAN DUTCH	NETHERLANDIC DUTCH
1. **schrikken 'to be startled' (465)**	1. **schrikken 'to be startled' (223)**
2. **werken 'to work' (346)**	2. **zich ergeren 'to be annoyed' (133)**
3. **zich ergeren 'to be annoyed' (232)**	3. **werken 'to work' (112)**
4. zich amuseren 'to enjoy oneself' (181)	4. zich schamen 'to be embarrassed' (80)
5. **betalen 'to pay' (167)**	5. **lachen 'to laugh' (70)**
6. **lopen 'to run' (125)**	6. **zich vervelen 'to be bored' (58)**
7. **lachen 'to laugh' (123)**	7. **lopen 'to run' (36)**
8. **zich vervelen 'to be bored' (78)**	8. **piekeren 'to worry' (22)**
9. rijden 'to ride' (69)	9. **betalen 'to pay' (18)**
10. **piekeren 'to worry' (65)**	10. zoeken 'to search' (14)

A notable difference is the presence in the left hand column of the verb *zich amuseren* 'to enjoy oneself', which only had 5 total occurrences in the Netherlandic Dutch data set. It appears that *zich amuseren* is more frequently used in Belgian Dutch newspapers in general: a search for [lemma="amuseren"] in the journalistic subcorpora of SoNaR yielded only 300 hits for Netherlandic Dutch versus almost 3,000 for Belgian Dutch. Even if we take into account the fact that the Belgian corpus is about three times larger than the Netherlandic corpus, the difference remains substantial. Another verb that is much more frequently intensified in Belgian Dutch is *rijden* 'to drive/ride' (with only 4 occurrences in Netherlandic Dutch). Again, the total frequency in the journalistic subcorpora is much higher in Belgian Dutch (approximately 65,500 hits for [lemma="rijden"]) than in Netherlandic Dutch (approximately 12,000 hits). The higher frequency of *rijden* 'to ride' in Belgian data may be due to a larger number of articles on various cycling competitions in Belgian newspapers.

(125) Minivoetbalclub 'De Bruine Beren' pakt elk zomer uit met een driedaagse 'Giro Del Vino', waarbij de deelnemers **zich de ziel uit het lijf rijden** om de gele trui te bemachtigen. (SoNaR-BE)
[...] the participants themselves the soul out of the body ride [...]
'Indoor soccer club 'De Bruine Beren' organises a three-day Giro Del Vino each summer, during which the participants ride their butts off to obtain the yellow jersey.'

The verb *zich schamen* 'to be embarrassed', by contrast, is proportionally much less frequently used in Belgian Dutch: with only 43 occurrences, it just falls outside of the top ten, whereas it was in 4th place with 80 tokens in the smaller Netherlandic Dutch data set (total frequency of [lemma="schamen"] in BE is 1,185 versus 1,204 in NL]. *Zoeken* 'to search', finally, was in 10th place in Netherlandic Dutch, whereas it is in 11th place in Belgian Dutch, so there is no big difference there. Turning our attention to the least frequent verbs, we see that the proportion of one-offs is only slightly lower in Belgian Dutch: there are 88/167 hapaxes in Belgian Dutch (53%) and 85/137 hapaxes in

Netherlandic Dutch (62%). The large proportion of hapaxes in Dutch in general testifies to the schematicity or openness of the verb slot of the intensifying fake reflexive resultative construction at the highest level of abstraction.

If we compare the upper panels of the Figures, the proportion of reflexive versus non-reflexive and transitive versus intransitive verbs is very similar in Belgian and Netherlandic Dutch. In both national varieties of Dutch, the majority of the data contain a non-reflexive intransitive verb (2,177 tokens in the entire SoNaR data set).

Table 4.3. Verb proportion (reflexivity and transitivity combined) in the intensifying set in SoNaR-BE and SoNaR-NL

	TOKENS (RES)	TYPES	TOKENS (RES)	TYPES
	BELGIAN DUTCH		NETHERLANDIC DUTCH	
Transitive verbs	350 **(+0.99)**	42	136 **(-0.99)**	39
Intransitive non-reflexive verbs	1,553 **(+2.03)**	119	624 **(-2.03)**	90
Intransitive reflexive verbs	542 **(-3.11)**	6	282 **(+3.11)**	8
TOTAL	2,445	167	1,042	137

A chi-square test on Table 4.3 shows that there was an overall significant difference with respect to the token frequencies of the verbal categories in Belgian and Netherlandic Dutch, but the effect size is small (χ^2=9.79, p=.007, df=2, V=.053). As we are dealing with more than two categories, the chi-square does not tell us whether all or, if not all, which of the categories show a significant effect. We can use the adjusted standardised residuals as a post-hoc test: generally, values exceeding +/-1.96 (or +/-2.0 by convention) indicate that the frequency in that cell is significantly higher/lower than would be expected at the statistical significance level of p=.05 (see, e.g., Agresti 2002, Sharpe 2015 on the use of standardised residuals as post-hoc correction).[37] As we are performing multiple comparisons, we could adjust our significance level for multiple testing (after Bonferroni correction, the new significance level is at p=.008, corresponding to a new critical z-value of +/-2.63), but it is customary to just use the standard of +/-2. The residuals are presented in bold in Table 4.3. It appears that the main source of the significant effect is in the reflexivity of the verb: the non-reflexive intransitive verbs are significantly more frequent in Belgian Dutch (residual of +2.03), whereas the reflexive intransitive verbs are significantly less frequent (residual of -3.11) than in Netherlandic Dutch.

Although the reflexive verbs are proportionally less token frequent in Belgian Dutch, a comparison of token and type frequencies again reveals their important relative contribution to the Belgian data set. In Belgian Dutch, the 542 reflexive tokens are distributed over 6 verbs, with 4 of these verbs accounting for approximately 99% of the

[37] We use the adjusted standardised residuals because they take into account the overall size of the sample when measuring the distance between the observed and expected counts.

reflexive tokens or 21% of the entire data set. As expected, there are some inherently reflexive verbs featured in the top ten of the most frequent verbs overall, viz. *zich ergeren* 'to be annoyed', *zich amuseren* 'to enjoy oneself' and *zich vervelen* 'to be bored'. If we look at the semantics of all verbs that are used in the Belgian data, we find a wide variety of verb types from the same semantic categories that were identified on the basis of the Netherlandic Dutch data.

Moving down to the lower left panel, we can be very brief about the use of the reflexive pronouns: the proportion of the different reflexive pronouns is – give or take one percent – exactly the same in both national varieties of Dutch.

To summarise, a close inspection of the verb slot and the reflexive pronoun does not reveal any notable differences between the two national varieties of Dutch, with the exception of the frequency discrepancy in the category of reflexive versus non-reflexive intransitive verbs. The intensifying fake reflexive resultative construction can be used to intensify a wide variety of verbal activities in Dutch, with 235 verb types overall, 133 of which are hapaxes. In constructional terms, the verb slot of the construction is highly schematic, although not *maximally* schematic so that any Dutch verb could fill it: the verbal activity must involve some aspect that can be boosted. There are 69 verbs which occur in the construction in both national varieties, and the lexical overlap is especially substantial in the category of highly frequent verbs. This indicates that some activities are inherently more susceptible to being intensified by this construction than others. In present-day Dutch, two broad categories can be distinguished: on the one hand, there are some very frequent verbs expressing a kind of emotional or cognitive experience, on the other, there is a large category of verbs denoting some kind of activity (physical activity, communicative activity, consumption activity...), some of which are also individually very frequent. Even though a large part of the data set does contain such (personal) "experience" verbs, the first person pronouns are less frequent than the third person pronouns. We have proposed that the specific nature and goals of the journalistic genre at least partly explain the dominance of third person clauses.

4.1.2.2 Intensifier

(a) General: national preferences

The Netherlandic data set featured 68 different intensifier types, 23 of which were hapaxes. In comparison, the Belgian data set contains 96 different intensifier types, 46 of which are hapaxes. Looking at the intensifier type-token ratios, we would have to conclude that Belgian Dutch has a slightly lower variety of different intensifiers (TTR=0.04) than Netherlandic Dutch (TTR=0.07). At the same time, almost half of the intensifier types in Belgian Dutch are hapaxes whereas the hapaxes take up "only" about a third of the intensifier types in Netherlandic Dutch. It has been assumed that "creative uses of a construction are adequately operationalised as hapax occurrences in a large

corpus" (Zeschel 2012: 185, 228), so this would indicate that the construction is actually used more creatively in Belgian Dutch than in Netherlandic Dutch. We will return to these creative hapaxes in section (c) below. Comparing the top ten of the most frequently used intensifiers in Belgian and Netherlandic Dutch, we see that there is an overlap of 7 intensifiers.

(xii) Top 10 intensifiers in Belgian and Netherlandic Dutch

BELGIAN DUTCH	NETHERLANDIC DUTCH
1. **te pletter 'to smithereens' (445)**	1. **rot 'rotten' (154)**
2. **blauw 'blue' (307)**	2. **suf 'drowsy' (123)**
3. **rot 'rotten' (241)**	3. **dood 'dead' (116)**
4. een hoedje 'a little hat' (200)	4. **kapot 'broken' (104)**
5. uit de naad 'out of the seam' (192)	5. **een ongeluk 'an accident' (60)**
6. **dood 'dead' (177)**	6. **te pletter 'to smithereens' (60)**
7. **suf 'drowsy' (129)**	7. groen en geel 'green and yellow' (44)
8. de ziel uit het lijf 'the soul out of the body' (98)	8. een slag in de rondte 'a punch around' (37)
9. **kapot 'broken' (72)**	9. wild 'wild' (34)
10. **een ongeluk 'an accident' (57)**	10. **blauw 'blue' (28)**

In the entire data set, there are 41 intensifiers that are used in both national varieties (see Appendix IV-1). We used the chi-square goodness-of-fit test to check whether the intensifiers showed a significant preference (at the statistical confidence level of p=.05) for either Belgian or Netherlandic Dutch, i.e. whether the observed frequencies in one of the two varieties are higher than the expected frequencies, if the total proportion of the data sets are taken into account (2,445 Belgian examples, 1,042 Netherlandic examples). Table 4.4 lists the overlapping intensifiers which were found to be significantly more frequent in one of the two national varieties of Dutch.

Table 4.4. Intensifiers with a significant preference for Belgian or Netherlandic Dutch

	BELGIAN DUTCH	NETHERLANDIC DUTCH	CHI-SQUARE
PREFERENCE FOR BELGIAN DUTCH			
blauw	307/2,445	28/1,042	χ^2=73.05;p<.0001
de pleuris	39/2,445	3/1,042	χ^2=9.31; p=.0023
een bult	53/2,445	1/1,042	χ^2=18.94; p<.0001
een hoedje	200/2,445	28/1,042	χ^2=32.87; p<.0001
krom	27/2,445	3/1,042	χ^2=4.75; p=.029
te pletter	445/2,445	60/1,042	χ^2=77.24; p<.0001
uit de naad	192/2,445	21/1,042	χ^2=39.81; p<.0001

	BELGIAN DUTCH	NETHERLANDIC DUTCH	CHI-SQUARE
PREFERENCE FOR NETHERLANDIC DUTCH			
dood	177/2,445	116/1,042	χ^2=12.72; p=.0004
een ongeluk	57/2,445	60/1,042	χ^2=24.56; p<.0001
groen en geel	3/2,445	44/1,042	χ^2=88.1; p<.0001
kapot	72/2,445	104/1,042	χ^2=70.27; p<.0001
rot	241/2,445	154/1,042	χ^2=15.2; p<.0001
suf	129/2,445	123/1,042	χ^2=42.19; p<.0001
wezenloos	2/2,445	22/1,042	χ^2=40.83; p<.0001

Of the 41 overlapping intensifiers, only 14 intensifiers display a significant effect: 7 intensifiers are significantly more frequent in Belgian Dutch, 7 intensifiers are preferred by speakers of Netherlandic Dutch. Most of the intensifiers that are found in only one of the two national varieties are low-frequency intensifiers or creative one-offs, as in examples (126) to (128) below.

(126) Nu moeten ze plots systemen lopen, maar ze **werken zich de ziel uit de naad**. (SoNaR-BE)
[...] but they work themselves the soul out of the seam
'Now suddenly they have to run systems, but they work very hard.'

(127) Toch **werkt** een meerderheid van ons **zich te sappel**, rennen we mee in de ratrace. (SoNaR-BE)
yet works a majority of us themselves to worries (sappelen = 'to worry') [...]
'Yet, most of us work very hard and participate in the rat race.'

(128) Grandaddy **toerde zich een eind in de rondte** en begon uiteindelijk zelfs de bloedeloze thuisstad Modesto, de barbecues en de vrienden daar te missen. (SoNaR-BE)
grandaddy toured itself a distance around [...]
'Grandaddy went on tour after tour and eventually even started missing the bloodless hometown Modesto, its barbecues and the friends over there.'

These are likely examples of idiosyncrasies at the level of the individual speaker and, in that regard, cannot really reveal any underlying differences with respect to the intensifier preferences between the two national varieties.[38] Still, there are a couple of hapax intensifiers which can be linked to national variation in a more general sense. For example, it is of course no coincidence that the intensifier *VLD-blauw* 'VLD-blue' is found in the Belgian data set, as the VLD (Flemish Liberals and Democrats) is a Belgian political party. In addition, the Belgian Dutch data set contains some typically Belgian variations

[38] Of course it is possible that *de ziel uit de naad* 'the soul out of the seam' is modelled on the Belgian exclusive intensifiers *de ziel uit het lijf* 'the soul out of the body' and *de naad uit het lijf* 'the seam out of the body'. It is a bit surprising, then, that *een eind in de rondte* 'a distance around' is found in the Belgian data, given that the similar intensifier *een slag in de rondte* 'a punch around' is exclusive to Netherlandic Dutch.

on the intensifier *het vuur uit de sloffen* 'the fire out of the slippers', replacing *sloffen* by *sloefen* or *slofkens*, which are colloquial Flemish terms for slippers.

More revealing of national preferences are the limited number of intensifiers that are relatively frequent (i.e. more than ten individual attestations) in one variety, but completely absent in the other. For Netherlandic Dutch, these are *een slag/slagen in de rondte* 'a punch/punches around' (37 singular, 9 plural attestations), *wild* 'wild' (34 attestations) and *ongans* 'unwell' (11 attestations).

(129) Beckham **knokte zich**, net als zijn ploeggenoten **een slag in de rondte**. (SoNaR-NL)
beckham fought himself, just like his teammates a punch around
'Beckham fought hard, just like his teammates.'

(130) Er zijn collega's die **zich drie slagen in de rondte werken** alvorens af te knappen. (SoNaR-NL)
there are colleagues who themselves three punches around work [...]
'There are colleagues who work extremely hard before having a breakdown.'

(131) De Hongkongse overheid was **zich wild geschrokken** van de rechterlijke uitspraak. (SoNaR-NL)
the hongkong government was itself wild startled [...]
'The Hongkong government was very startled by the judicial decision.'

(132) New York houdt de adem in voor aanslagen die zeker moeten komen - en **koopt zich** intussen **ongans**. (SoNaR-NL)
[...] and buys itself in the meantime unwell
'New York holds its breath for the attacks that are sure to come – and in the meantime, people buy everything their heart desires.'

For Belgian Dutch, we find *de ziel uit het lijf* 'the soul out of the body' (98 attestations), *een aap* 'a monkey' (32 attestations), *steendood* 'stone-dead' (16 attestations), *een beroerte* 'a stroke' (13 attestations), *zot* 'crazy' (13 attestations) and *de naad uit het lijf* 'the seam out of the body' (12 attestations).

(133) Ik **huil me de ziel uit het lijf**, van geluk. (SoNaR-BE)
I cry myself the soul out of the body of happiness
'I am crying like a baby with joy.'

(134) De beleggers **schrokken zich een aap** en de beurskoers kelderde. (SoNaR-BE)
the investors startled themselves a monkey [...]
'The investors were very startled and the stocks plummeted.'

(135) Ik begrijp de mensen niet die daar een volledige maand hun vakantie kunnen doorbrengen: volgens mij **verveel** je **je steendood** na drie dagen. (SoNaR-BE)
[...] bore you yourself stone-dead after three days
'I don't understand that people can spend an entire month on holiday there: I think you're bored out of your mind after three days.'

(136) Uw dienaar **schrok zich een beroerte** toen naar hem werd gewezen. (SoNaR-BE)
your servant startled himself a stroke [...]
'Your servant was very startled when he was suddenly pointed at.'

(137) Elk jaar **sparen** studenten van over de hele wereld **zich zot** om een plek te veroveren in onze prestigieuze modeopleidingen. (SoNaR-BE)
every year save students from over the entire world themselves crazy [...]
'Every year, students from the entire world save up a lot of money to conquer a spot in one of our prestigious fashion educations.'

(138) De technici **werkten zich de naad uit het lijf** om het verkeer tegen de avondspits weer normaal te kunnen doen verlopen. (SoNaR-BE)
the technicians worked themselves the seam out of the body [...]
'The technicians worked very hard to get the traffic back to running smoothly by the evening rush-hour.'

While some of these lexical items may be somewhat regionally marked (e.g. *zot* is the preferred term for 'crazy' in (informal) Belgian Dutch, whereas *gek* is more often used in Netherlandic Dutch and *steendood* 'stone-dead' is labelled as Belgian Dutch in Van Dale), most of the individual lexemes (i.e. *de ziel, wild, het lijf, de naad, een aap* etc.) do not appear to be typically Belgian or Netherlandic words per se. It would seem that most of these words or phrases are "nationally exclusive" only when they are used in this specific combination, i.e. as an intensifier in this particular construction. Together with the results in Table 4.4, this confirms the hypothesis that speakers of Belgian and Netherlandic Dutch do hold somewhat different preferences with respect to the intensifiers they use in this construction. In the remainder of this paragraph, we will have a closer look at the syntactic and semantic properties of these intensifiers to see whether they disclose any other prominent differences between Belgian and Netherlandic Dutch.

(b) Syntactic categories

If we compare the bottom right panels of the circle diagrams, we already see some subtle shifts in the way in which the syntactic categories are distributed. If we have another look at Table 4.4, we notice that most of the intensifiers with a preference for Netherlandic Dutch are adjectives, whereas the intensifiers that show a preference for Belgian Dutch are syntactically more diverse. In addition, the Belgian Dutch data contain one NP+particle intensifier, *het licht uit* 'the light out', a category which was absent in Netherlandic Dutch.

(139) Dan binden ze zes dagen op zeven, tot vier uur per dag, de schaatsen om en **trainen** ze **zich het licht uit**. (SoNaR-BE)
[...] and train they themselves the light out
'Then they put on their skates six days out of seven, up to four hours a day, and train their butts off.'

Table 4.5 presents a more detailed picture of the frequency distributions of the syntactic categories.

Table 4.5. Frequency comparison of syntactic category of the intensifier in SoNaR

	Tokens (res)	Types	Int. TTR	Hapax types	Verb range
Belgian Dutch					
AP	1,062 **(-12.16)**	35	0.033	14	97
NP	458 **(+0.41)**	26	0.057	15	45
PP	671 **(+11.69)**	5	0.007	1	81
NP+PP	252 **(+3.73)**	27	0.107	14	37
NP+AP	1 **(-0.62)**	1	1	1	1
NP+part	1 **(+0.65)**	1	1	1	1
Total	2,445	96	0.039	46	167
Netherlandic Dutch					
AP	687 **(+12.16)**	25	0.036	6	97
NP	189 **(-0.41)**	24	0.127	11	50
PP	99 **(-11.69)**	4	0.040	0	28
NP+PP	66 **(-3.73)**	14	0.212	5	25
NP+AP	1 **(+0.62)**	1	1	1	1
NP+part	0 **(-0.65)**	0	0	0	0
Total	1,042	68	0.064	23	137

The token frequency differences were found to be highly significant overall with a medium effect size (χ^2=193.83, p<.0001, df=5, V=.236), but the standardised residuals in bold show that not all syntactic categories display significant differences in the national varieties of Dutch.[39] Even so, the similarities or differences may not just be situated at the token frequency level but may also concern type frequency, type-token ratio or verbal range. In terms of relative token frequency, the nominal category is about as frequent in Belgian Dutch as in Netherlandic Dutch, but there are proportionally more different nominal types in Netherlandic Dutch. Again, however, the hapaxes are slightly better represented in Belgian Dutch, with 57% versus 46% nominal hapaxes. The NP+PP category is better represented in Belgian Dutch in terms of total frequency (residual of +3.73) and it scores remarkably higher in terms of hapax types (cf. infra). The most salient differences between Belgian and Netherlandic Dutch concern the prepositional and adjectival intensifiers. Although the prepositional category accounts for a much larger portion of the data in Belgian Dutch than in Netherlandic Dutch (residual of +11.69), the difference in type frequency is minimal. If we look at the top ten intensifiers in Belgian Dutch in (xii), the reason for the large discrepancy in token frequency becomes clear. The most frequent intensifier in Belgian Dutch is *te pletter* 'to smithereens', already accounting for 445 tokens (i.e. 66% of the prepositional tokens and 18% of the entire data set); if we

[39] Again, we could apply a Bonferroni-correction (adjusted significance level p=.004, adjusted z-value +/-2.87) but it does not affect the overall results.

add *uit de naad* 'out of the seam', we arrive at 95% of the prepositional set and 26% of the entire data set. The PP intensifiers in Belgian Dutch also show signs of a much wider coverage in terms of verbal range than in Netherlandic Dutch. As we will see in §4.2.2 and §4.3.2, not only is *te pletter* the most frequent intensifier, it also has the highest degree of combinatorial flexibility in Belgian Dutch. The larger contribution of prepositional intensifiers entails that the adjectival intensifiers are significantly less dominant in Belgian Dutch (residual of -12.16). Aside from the difference in relative token frequency, however, the behaviour of the adjectival categories is largely similar in Belgian and Netherlandic Dutch: the proportion of types and hapaxes, as well as the verbal range is about the same in both national varieties. Teasing apart the reflexive and non-reflexive verbs in Figure 4.4 reveals that the adjectival intensifiers are the preferred category for the reflexive verbs, just like in Netherlandic Dutch.

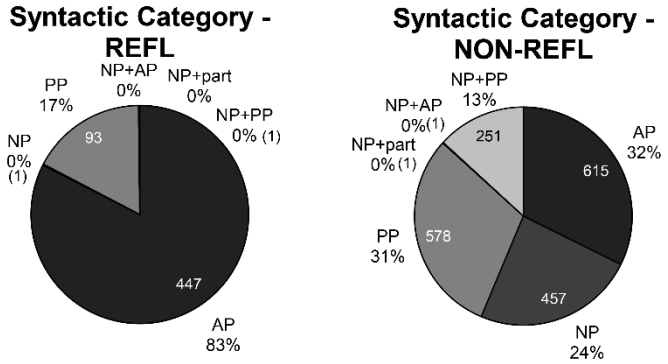


Figure 4.4. Comparison of the syntactic category of intensifiers combining with reflexive versus non-reflexive verbs in SoNaR-BE

(c) Semantic categories

The intensifiers in the Belgian Dutch data set seem to be drawn from the same semantic or conceptual domains as the Netherlandic intensifiers for the most part. In Belgian Dutch, as well, category I of the negatively connoted states is extremely well-represented in terms of total number of intensifiers. Some new adjectival intensifiers that were not in the Netherlandic data set are listed below.

I. Negatively connoted states: *gaar* 'cooked', *kreupel* 'crippled', *murw* 'mellow', *onnozel* 'silly', *zot* 'crazy'...

Based on the quote by Cappelle (2014: 273) that "many of the intensifications are not common in Flanders (the Dutch-speaking part of Belgium) at all. This is especially true for patterns with diseases", we would have expected the second category of the diseases to be much less important in Belgian Dutch. However, some of the same diseases that were found in the Netherlandic Dutch data – and one previously unattested disease (*de typhus* 'the typhoid') – occasionally popped up in the Belgian data as well. Granted, however, the variety of diseases is slightly smaller. Most of these intensifiers are not very

token frequent (just like in Netherlandic Dutch), with the exception of the intensifier *de pleuris* 'the pleurisy' (39 tokens in Belgian Dutch, versus only 3 in Netherlandic Dutch).

II. Diseases: *apelazerus* 'fictitious disease', *de pleuris* 'the pleurisy', *de tering* 'the consumption', *de typhus* 'the typhoid', *een delirium* 'a delirium', *het apezuur* 'fictitious disease', *het lazerus* 'the leprosy', *het leplazerus* 'fictitious disease'...

The inalienable possession category is also well-represented in Belgian Dutch. It contains several one-off variations of higher frequency intensifiers. For example, we do not only find *het vuur uit de sloffen* 'the fire out of the slippers' but also *uit de sokken* 'out of the socks', *uit de sloefen* 'out of the slippers (regional term)' and *uit de slofkens* 'out of the little slippers (regional term)'. In addition, *de ziel* 'the soul' and *de naad* 'the seam' are popular NPs that can be used with multiple different PP complements that were absent in Netherlandic Dutch.

III. Inalienable possession: *de naad uit de broek* 'the seam out of the pants', *de naad uit het lijf* 'the seam out of the body', *de ziel uit het lijf* 'the soul out of the body', *de ziel uit de naad* 'the soul out of the seam', *de ziel uit de raap* 'the soul out of the head', *het vuur uit de sloffen/sloefen/slofkens* 'the fire out of the slippers'...

Next, there are some new colour terms in Belgian Dutch that add strength to the hypothesis that intensifiers can outclass other intensifiers in terms of booster strength (cf. supra). For example, *donkerblauw* 'dark blue' denotes a darker shade than *blauw* 'blue' and could therefore be interpreted as a stronger intensifier. Other examples involve a more *specific* characterisation of the colour shade and, in that regard, could also be considered as potentially stronger intensifiers, e.g. *adellijk blauw* 'noble blue',*VLD-blauw* 'VLD-blue', *bicblauw* 'pen blue' and *spinaziegroen* 'spinach green'. In any case, the creative (often context-specific) colour variations add some extra flavour to the discourse and are more likely to attract the reader's attention than the conventional colour terms.

IV. Colour: *adellijk blauw* 'noble blue', *bicblauw* 'pen blue', *blauw en paars* 'blue and purple', *donkerblauw* 'dark blue', *VLD-blauw* 'VLD-blue', *groen* 'green', *spinaziegroen* 'spinach green', *het geel* 'the yellow'...

Finally, there are also some intensifiers that do not immediately fit into any of the above-mentioned semantic categories.

V. Other: *een aap* 'a monkey', *een pissebed* 'an isopod', *beten en scheten* 'bites and farts', *het licht uit* 'the light out'...

The hapaxes in Belgian Dutch behave in much the same way as they do in Netherlandic Dutch. Most of the hapaxes belong to one of the established semantic domains; as we pointed out, there are some variations on more frequent intensifiers, which have likely

been created as analogical extensions on the basis of a more frequent model, see (140) and (141).

(140) Al maanden **loopt** mijnheerke Louis **zich het vuur uit zijn slofkens** om iedereen te vertellen dat hij, in tegenstelling tot den Daems, geboren, getogen en zelfs gedoopt is in Leuven. (SoNaR-BE)
for months runs mister louis himself the fire out of his little slippers [...]
'For months, mister Louis has been running around telling everyone that he, unlike Daems, is born and bred in Leuven.'

(141) Tijdens de Boerenkrijg **ergerden** de sansculotten **zich spinaziegroen** omdat zij de hoeve niet konden vinden. (SoNaR-BE)
[...] annoyed the sansculottes themselves spinach-green [...]
'During the Peasant's Revolt the sansculottes were very annoyed because they couldn't find the farm.'

Occasionally, we also find a truly creative invention of the speaker, as in (142) below.

(142) Ik **lach mij** soms **beten en scheten** met iets, maar probeer dat de volgende dag te reconstrueren en er is niets meer. (SoNaR-BE)
I laugh myself sometimes bites and farts [...]
'Sometimes I'm laughing my head off over something, but when I try to reconstruct it the next day, I can't.'

In the next section, we will move on from the lexical fillers of the individual slots to the interaction between the different slots of the construction by studying the collocational patterns in the intensifying fake reflexive resultative construction.

4.2 Collocational patterns

The previous section revealed some interesting discrepancies between the general use of the construction in Belgian and in Netherlandic Dutch. While the instantiation of the verb slot was remarkably parallel in both national varieties, there were some clear differences with respect to the intensifier slot. If we bring together the frequency information of all the individual elements of the construction from the previous paragraph, we could construct the "prototypical" instance of the intensifying fake reflexive resultative construction as a sentence in which we have the most frequent verb, the most frequent intensifier and the most frequent form of the reflexive pronoun, i.e. third person singular. For Netherlandic Dutch, we get *zich*$_{\text{sing}}$ *rot schrikken* 'to startle oneself rotten'. With 26 instances, this exact combination is indeed the most frequently occurring configuration in the data set – i.e. the construct type with the highest number of empirically attested

construct tokens, cf. Zeschel (2012: 15). For Belgian Dutch we would get *zich*$_{sing}$ *te pletter schrikken* 'to startle oneself to smithereens', but this is *not* the most frequent configuration of individual elements at all. The 9 tokens of the specific configuration *zich*$_{sing}$ *te pletter schrikken* pale in comparison with the 122 occurrences of *zich*$_{sing}$ *een hoedje schrikken* 'to startle oneself a little hat' or the 82 instances of *zich*$_{sing}$ *uit de naad werken* 'to work oneself out of the seam'. Even *zich*$_{sing}$ *suf piekeren* 'to worry oneself drowsy', which combines the 10th most common verb and the 7th most common intensifier, respectively, has more occurrences than the combination of the most frequent verb and the most frequent intensifier. From a constructional point of view an approach in which we merely combine the most frequent individual elements is too naive. In this section, we will argue that the way in which verbs and intensifiers are combined is not solely determined by their individual token frequencies but can be influenced by a number of other factors. If we want to draw an accurate picture of the use of the intensifying fake reflexive resultative construction in present-day Dutch, we need to look at the interrelationship between verbs and intensifiers and how it is influenced by both convention and linguistic creativity.

4.2.1 Synchronic use

In section 4.1 we have identified semantic groups of verbs and intensifiers on the basis of introspection and basic world knowledge. It is often assumed in distributional semantic theories that items which form a semantic cluster exhibit the same collocational behaviour. The hierarchical cluster method (City-block distance measure and Ward's method) is generally used to cluster together a set of node words on the basis of their shared collocates within a certain range, thus providing a more objective way of identifying semantic classes of verbs that are used in a construction. Gries & Stefanowitsch (2010) first applied a cluster analysis to the ditransitive construction (i.e. the matrix verb and its collocates in the same sentence) to see whether the subsenses distinguished by Goldberg (1995) (e.g. implied transfer, enabling of transfer, communication, etc.) are identifiable as separate clusters. Some of the clusters that are returned by the analysis are partly semantically interpretable, but they are generally incomplete or contain some verbs that should not belong in that cluster. For example, the formal communication verbs like *assure*, *persuade* or *inform* are clustered together, but *convince* is not part of that cluster. Instead, the cluster of formal communication verbs contains a verb like *lend*, which is not a communication verb at all. These results led them to suggest a particular application of the method that is not concerned with the *linear* collocates of individual items, but with the covarying collexemes, i.e. items that occur together in two slots within the same construction, but are not necessarily immediately adjacent (Stefanowitsch & Gries 2005) – cf. infra, where we will also apply a covarying

collexeme analysis. For example, Gries & Stefanowitsch (2010) have shown that clustering the words in one slot according to the words in another slot yields clusters that form more or less coherent semantic classes for the *into*-causative and the *way*-construction. In the *into*-causative, the matrix verbs are grouped together on the basis of their covarying result gerunds (V_1 into V_2-*ing*). The analysis gives relatively semantically homogeneous clusters such as physical force verbs (e.g. *coerce*, *force*...), positive-persuasion verbs (e.g. *entice, tempt*...), negative-persuasion verbs (*embarrass, panic, shame*...) and trickery verbs (*dupe, fool, trick*...). In the *way*-construction, then, it was investigated whether it is possible to distinguish semantic verb classes by looking at the covariation of verbs with specific prepositions (V POSS *way through/into/to/over*...). It appears that the clusters again correspond to relatively coherent groups, e.g. a small cluster of physical force verbs (*force, work*...) and a super-cluster with slow-movement verbs (e.g. *thread, wend, wind*, *worm*...). However, the results of this "special" application of the cluster analysis are not always as unequivocal. Zeschel (2012: 218-227) applied the (collexeme) cluster analysis to different perception intensity collocations in English and German (e.g. Int + N: *glowing health* and *sirrende Hitze* 'buzzing heat'; Int + ADJ: *blisteringly fast* and *knackig kalt* 'lit. cracky cold'; Int + *with/vor* + V *to seethe with anger* and *kochen vor Begeisterung* 'boil with enthusiasm') and found that, especially in English, many near-synonyms or semantically related words do not always end up in the same cluster. He suggests that individual items often retain idiosyncratic preferences that are not shared by semantically similar items. In spite of the sometimes unexpected groupings, Zeschel's clusters do reveal interesting semantic or conceptual patterns in the intensity collocations as well. For example, the analysis provides support for the metaphorical association between heat and strong emotions or light and joy: words like *heat* and *pain* are often associated with the same type of intensifiers, as are, e.g., *light* and *smile*.

As a first exploration of a possible semantic basis for the collocational patterns in the intensifying fake reflexive resultative construction, we also applied the hierarchical cluster analysis to the covarying verbs and intensifiers in the construction. To keep the dendrograms somewhat neat, we only investigated the top twenty verbs (i.e. all attested combinations of these verbs with all the intensifiers in the data set), and the top twenty intensifiers (i.e. all attested combinations of these intensifiers with all the verbs in the data set). In order to ensure comparability with Belgian Dutch in the next subsection, we selected the top twenty verbs and the top twenty intensifiers in the entire SoNaR data set (i.e. SoNaR-NL and SoNaR-BE combined), only taking into account the *overlapping* items that had at least one occurrence in both data sets.

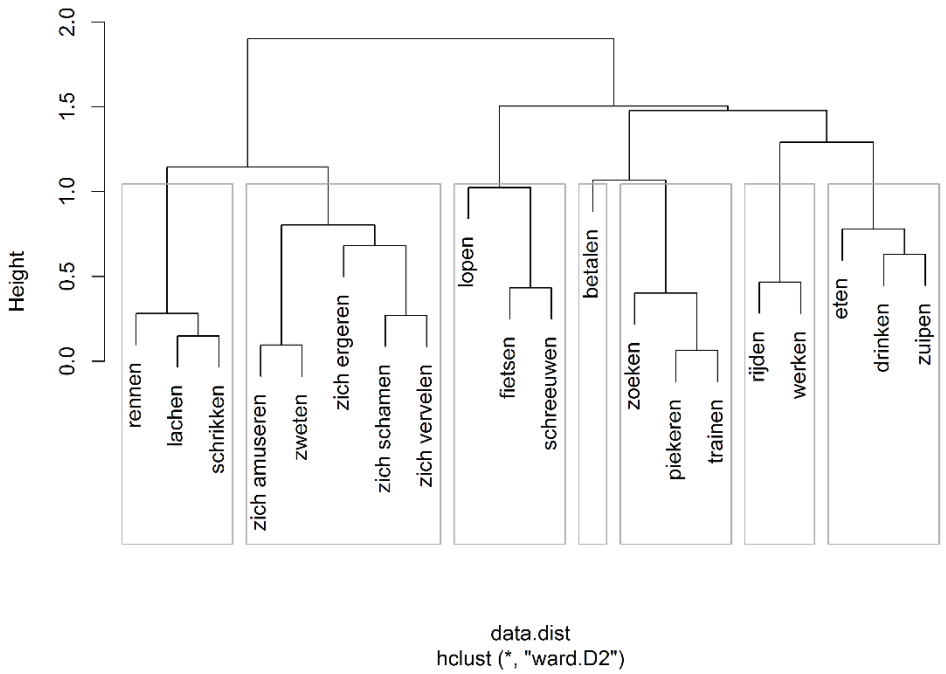


Figure 4.5. Dendrogram HCA of the top 20 verbs in SoNaR (clustered according to covarying intensifiers)

In Figure 4.5, the rightmost cluster is intuitively meaningful and clearly corresponds to one of the verbal semantic classes that were identified earlier, viz. the verbs of consumption. The second cluster from the right is also semantically homogeneous in as far as it contains two physical activity verbs, but other physical verbs are scattered across the remaining clusters (e.g. *rennen* 'to run' in the first cluster from the left, *lopen* 'to run' and *fietsen* 'to cycle' in the third cluster from the left and *trainen* in the fifth cluster from the left). Most clusters are heterogeneous in that they contain at least one "odd duck", e.g. *zweten* 'to sweat' in the second cluster from the left and *schreeuwen* 'to scream' in the third cluster from the left. Aside from the presence of *zweten* 'to sweat', the second cluster is interesting because it groups together the four inherently reflexive verbs, suggesting that these verbs do exhibit similar behaviour in the construction. As was pointed out before, the reflexive verbs also share a certain meaning component, so it is not entirely clear whether the shared collocational patterns are to be attributed to their semantics or to their inherent reflexivity, or both.

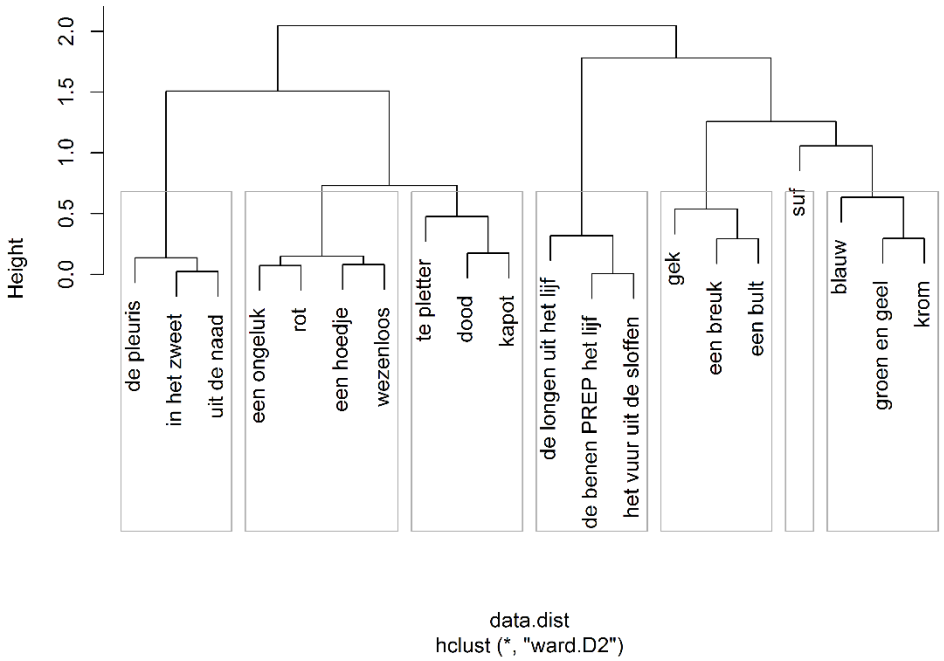


Figure 4.6. Dendrogram HCA of the top 20 intensifiers in SoNaR (clustered according to covarying verbs)

In Figure 4.6, we also recognise some of the semantic classes distinguished in §4.1.1.3 above. Some of the members of the inalienable possession or removal category are grouped together in the fourth cluster from the left and the group of *te pletter* 'to smithereens', *dood* 'dead' and *kapot* 'broken' is also semantically or conceptually coherent in that all of these can be used to refer to death. Other clusters generally contain a couple of intensifiers that could be expected to cluster together based on their semantics (e.g. the colour adjectives in the rightmost cluster and two of the less prototypical diseases in the fifth cluster from the left), but they are overall rather semantically heterogeneous. A possible explanation for why *suf* 'drowsy' forms a cluster of its own, rather than pairing up with the other negatively connoted adjectives like *rot* 'rotten', *wezenloos* 'blank/vacant' and *gek* 'crazy', will be given later in this section.

The results for the intensifying fake reflexive resultative construction are very similar to those for the ditransitive construction (Gries & Stefanowitsch 2010) and the intensity collocations (Zeschel 2012): some semantic patterns can be discerned, but the clusters are often incomplete or only partly homogeneous. Clearly, depending on the criterion that is used to define verb and intensifier classes (intuition on lexical semantics or collocational behaviour), we get somewhat different results: although some of the clusters in Figure 4.5 and Figure 4.6 do correspond to the semantic classes that were identified on an intuitive basis earlier, shared semantics fail to explain some of the collocational behaviour of verbs and intensifiers in this particular construction. The fact that semantically similar items may end up in different clusters could indicate that some individual verbs and/or

intensifiers have undergone collocational specialisation whereas their related items have not – at least, this is what Zeschel (2012: 218) proposes for the unexpected clusters in his data. Conversely, intensifiers which have very different lexical semantics can still pair up with the same verbs and verbs need not be semantically related to pair up with the same intensifier. A possible explanation is that for some intensifiers, the semantic bleaching is so advanced that the original lexical meaning no longer imposes any restrictions on their collocational range (cf. infra, §4.3).

Let us now look at these collocations in somewhat more detail. By just eyeballing the data, we already noticed that some verb-intensifier combinations occurred with a much higher frequency than others. Before we subject the data to empirical measures of productivity in the next section, we already want to have a first look at the verb-intensifier co-occurrences that are attested in the construction. A method that has proved successful in giving a first indication of the potential interaction of two slots within the same construction is the covarying collexeme analysis, one of the collostruction methods developed by Anatol Stefanowitsch and Stefan Gries (Stefanowitsch & Gries 2003, 2005, Gries & Stefanowitsch 2004a, R-script by Gries 2007). The covarying collexeme analysis determines which collexeme combinations (i.e. verb-intensifier combinations) occur more frequently than expected given the overall frequencies with which the individual collexemes (i.e. verbs and intensifiers) occur in the construction. The default association measure is the Fisher-Yates exact significance test, which has the advantage of not making any distributional assumptions. The Fisher exact p-value incorporates the effect size and assigns more weight to associations that are based on a higher number of tokens. That is, an observed co-occurrence of 25% is ranked higher if it is based on 25 out of 100 tokens rather than 1 out of 4 (Gries 2012, 2015). The p-values are mainly used "as an indicator of relative importance" (Stefanowitsch & Gries 2003: 239), that is, as underlying values to rank the attracted (or repelled) collexemes according to their association strength. In the covarying collexeme analysis, the actual measure of collostruction strength is the (negative) log-transformed p-value, p_{log10}, which is more easily interpretable. Log-transformed values of 1.30103 and higher correspond to statistical significance at the level of 5%; the higher the collostruction value, the lower the probability of error.

Table 4.6. Top 20 attracted and top 10 repelled covarying collexemes in SoNaR-NL

VERB	INTENSIFIER	FREQ. V	FREQ. INT	OBS. V-INT	EXP. V-INT	$\Delta P_{VERB-TO-INT}$	$\Delta P_{INT-TO-VERB}$	COLL. STR.
zich ergeren	groen en geel	133	44	44	5.62	0.33	0.91	42.43
lopen	het vuur uit de sloffen	36	14	14	0.48	0.39	0.98	21.69
schrikken	een hoedje	223	28	27	5.99	0.12	0.77	17.29
piekeren	suf	22	123	20	2.6	0.81	0.16	16.91

betalen	blauw	18	28	12	0.48	0.65	0.42	15.79
schrikken	rot	223	154	74	32.96	0.23	0.31	15.21
zich schamen	dood	80	116	35	8.91	0.35	0.25	14.47
lopen	de benen PREP het lijf	36	11	10	0.38	0.28	0.88	14.16
praten	de blaren op de tong	9	8	6	0.07	0.66	0.75	11.87
werken	uit de naad	112	21	15	2.26	0.13	0.62	10.42
werken	in het zweet	112	10	10	1.07	0.09	0.9	9.85
zuipen	klem	14	9	6	0.12	0.43	0.66	9.85
zich ergeren	wild	133	34	19	4.34	0.13	0.45	8.96
zich vervelen	te pletter	58	60	17	3.34	0.25	0.24	8.54
peinzen	suf	8	123	8	0.94	0.89	0.07	7.51
eten	ongans	12	11	5	0.13	0.41	0.45	7.46
zich schamen	de ogen uit het hoofd	80	6	6	0.46	0.08	0.93	6.77
lachen	een kriek	70	5	5	0.34	0.07	0.94	5.92
drinken	een stuk in de kraag	11	4	3	0.04	0.27	0.74	5.46
zich vervelen	dood	58	116	19	6.46	0.23	0.12	5.36
schrikken	suf	223	123	1	26.32	-0.14	-0.23	12.19
zich ergeren	suf	133	123	2	15.7	-0.12	-0.13	5.42
werken	dood	112	116	2	12.47	-0.1	-0.1	3.94
werken	suf	112	123	3	13.22	-0.1	-0.09	3.51
werken	rot	112	154	6	16.55	-0.11	-0.08	3.02
zich ergeren	te pletter	133	60	1	7.66	-0.06	-0.12	2.65
lachen	suf	70	123	2	8.26	-0.1	-0.06	2.17
zich vervelen	suf	58	123	2	6.85	-0.09	-0.04	1.63
lachen	te pletter	70	60	1	4.03	-0.05	-0.05	1.12
lopen	rot	36	154	2	5.32	-0.1	-0.03	1.1

The results in Table 4.6 give us a first indication of the verb-intensifier combinations that occur much more frequently than we would expect by chance. The top collocations should sound fairly conventional, or at least familiar, to any native speaker of Netherlandic Dutch. It is worth pointing out that a high collostruction strength does not necessarily imply that there is *symmetric* attraction between verb and intensifier: the association is often influenced by the limited combinatorial flexibility of one of the two elements. The direction of the attraction is indicated by the ΔP-values, viz. $\Delta P_{\text{verb-to-int}}$ or

$\Delta P_{int\text{-}to\text{-}verb}$. If the attraction is heavily asymmetric, one of the two values is considerably higher than the other (a difference of ≥0.5, Gries 2013a). For example, *zich groen en geel ergeren* 'to annoy oneself green and yellow', *zich een hoedje schrikken* 'to startle oneself a little hat' and *zich in het zweet werken* 'to work oneself in the sweat' are returned as strong collocations because the intensifier (almost) exclusively associates itself with just the verbs in question, even though *zich ergeren* 'to be annoyed', *schrikken* 'to be startled' and *werken* 'to work' are all frequently found outside of this collocation as well. In contrast, the collocation *zich suf piekeren* 'to worry oneself drowsy' is largely attributable to the low degree of combinatorial flexibility of the verb *piekeren* 'to worry', while *suf* 'drowsy' has a wide range of uses outside of the collocation. In case of more mutual, symmetric attraction, both values are much closer together. Symmetric attraction does not necessarily imply that the verb and intensifier in question are almost exclusively associated with one another; both verb and intensifier may well have a range of uses outside of this particular collocation but still co-occur more frequently than would be expected based on their individual frequencies, which is the case for, e.g., *zich rot schrikken* 'to startle oneself rotten' or *zich te pletter vervelen* 'to bore oneself to smithereens'. Below the bolded line, we find the repelled collocations, i.e. verbs and intensifiers that co-occur less frequently than one would expect on the basis of their individual token frequencies in the data set (hence the negative ΔP-values). We have limited ourselves here to combinations that do have at least one attestation, but there are of course many more potential combinations that do not occur at all. There appears to be something going on with *suf* 'drowsy' here as well (cf. the cluster analysis supra): even though *suf* 'drowsy' is one of the most frequently used intensifiers, it rarely pairs up with some of the highly frequent verbs like *schrikken* 'to be startled', *zich ergeren* 'to be annoyed' and *werken* 'to work'. We will come back to this below.

As an exploratory analysis of collocational behaviour, the covarying collexeme analysis clearly performs rather well. At the same time, there is a lot of information that is not taken into account in the calculation of the collostruction strength and that may be important in explaining certain phenomena like statistical preemption or blocking, analogical model formation or large scale productivity. The covarying collexeme analysis is based on the token co-occurrence frequencies of verbs and intensifiers and the token frequencies of both elements outside of this co-occurrence, but it does not take into account the type frequency of either of the elements, which will be shown to play an important role in productivity (cf. infra). Neither does it take into account the distribution of all tokens across the different types: the tokens could be (more or less) evenly distributed across the types, or they could have a very skewed (Zipfian) distribution, in which case just a few types already account for most of the tokens. A measure that has been suggested to capture the type-token distribution is relative entropy H_{rel}. H_{rel} is a value between 0 and 1, in which 0 indicates that one type accounts for all tokens and 1 stands for a perfectly equal distribution across the different types

(Gries 2012, 2013b). A category with a skewed distribution (or low variance), indicated by a lower H_{rel} value, has been argued to be more predictable and therefore easier to learn (Goldberg et al. 2004). In the next section, we will discuss how the skewness of the distribution may also influence the extensibility of the construction. A more comprehensive overview of all this frequency information can be provided in a cross-table. As it would be impossible to fit all 137 verbs and 68 intensifiers into a surveyable cross-table, Table 4.7 only shows the combinations of the top 15 verbs and the top 15 intensifiers. The overall frequencies that are given for token, type and hapax count, as well as the relative entropy measure are based on the entire data set.

Table 4.7. Full cross-tabulation for the top 15 verbs/intensifiers in SoNaR-NL

	rot	suf	dood	kapot	een ongeluk	te pletter	groen en geel	een slag in de rondte	wild	blauw	een hoedje	wezenloos	uit de naad	het vuur uit de sloffen	de longen uit het lijf	Tokens	Types	Hapaxes	H_{rel}
schrikken	74	1	24	19	24	13			14		27	11				223	19	8	0.514
zich ergeren	14	2	20	16		1	44		19	9		3				133	14	6	0.469
werken	6	3	2	18	5	10		14					15			112	22	5	0.638
zich schamen	21		35	18												80	4	0	0.296
lachen	20	2	6	12	9	1					1					70	17	7	0.544
zich vervelen	5	2	19	9	2	17										58	9	2	0.408
lopen	2	2				1							1	13	3	36	10	5	0.416
piekeren	1	20								1						22	3	2	0.087
betalen		2								12						18	4	1	0.232
zoeken	3	4			2			2	1			2				14	6	1	0.407
zuipen					1			1								14	8	6	0.421
eten		1	2					1								12	7	5	0.404
drinken				1	1	1		1								11	9	8	0.499
trainen		6						1					1			11	6	5	0.338
praten		1														9	3	1	0.202
Tokens	154	123	116	104	60	60	44	37	34	28	28	22	21	13	12				
Types	11	61	13	16	23	22	1	22	3	7	2	9	6	1	8				
Hapaxes	5	38	4	9	18	18	0	19	1	3	1	6	4	0	6				
H_{rel}	0.353	0.751	0.383	0.441	0.47	0.473	0	0.516	0.161	0.297	0.031	0.341	0.212	0	0.393				

In Chapter 2, it was argued that there is an open schema [SUBJ V REFL INT], associated with the semantics 'Subj Vs excessively/intensely' (cf. Ch3, §3.3.6 for slight semantic variations), in which the verb slot and the intensifier slot have a high degree of "openness". Although the possibilities are countless in theory, some individual items present themselves as more eligible candidates than others: 65% of all the uses of the construction in present-day Dutch newspapers are already accounted for by just the intercombinations of the top 15 verbs and intensifiers. It was also suggested that there is a large degree of conventionality involved when it comes to combining verbs and intensifiers. Table 4.7 corroborates that there is some kind of covariation or interaction between the verb slot and the intensifier slot: it appears that the choice for a particular verb (or intensifier) immediately projects a range of possible fillers onto the other open slot. We could even state that all of these potential fillers are competing among each other to be selected as the "ultimate" slot filler and that some items appear to be more successful at this than others (cf. Ch2, §2.3 and Chapter 5 for a diachronic point of view).

Essentially, the cross-table can be read in two ways. First, we can start out from the perspective of the verbs and see which intensifiers they combine with on the horizontal axis. The two most frequent verbs, viz. *schrikken* 'to be startled' and *zich ergeren* 'to be annoyed', are also the most flexible verbs in this table, both occurring with 9 out of the top 15 intensifiers. However, if we take into account the entire data set, it is *werken* 'to work' that comes out on top with 22 different intensifier types. Judging by the rather high value of H_{rel}, the distribution of *werken* is much less dependent upon a number of very frequent types than the distribution of some of the other verbs (cf. the next section on how this relates to productivity). In contrast, the verb *zich schamen* 'to be embarrassed', despite its relatively high token frequency, is found with only 4 types and has a much lower value for relative entropy, indicating that the distribution is limited to a number of very frequent types. Second, we could take the perspective of the intensifiers and look at the co-occurring verbs on the vertical axis. The intensifiers *groen en geel* 'green and yellow' and *het vuur uit de sloffen* 'the fire out of the slippers' immediately jump to the eye: their relative entropy values are at 0 because they enter into an exclusive collocation with just one verb, viz. *zich ergeren* 'to be annoyed' and *lopen* 'to run', respectively. In fact, the reason that these two intensifiers are featured among the top intensifiers at all is their being part of a frequently used collocation. The importance of teasing apart type and token frequency is also confirmed by the fact that the most frequent intensifier *rot* 'rotten' is not the most flexible intensifier, occurring with 9 out of the top 15 verbs but with only 11 verbs overall. The strong preference for the verb *schrikken* 'to be startled' is reflected in the rather low H_{rel} value. The most "promiscuous" intensifier is *suf* 'drowsy', which covaries with 12 out of the top 15 intensifiers and no less than 61 verb types in the entire data set. The high H_{rel} value of 0.751 suggests that *suf* shows a markedly balanced distribution across its different types. If we have another look at the combinations with

the top 12 verbs, we generally find remarkably low values. The fact that the attested combinations with the most frequent verbs are so infrequent may be the reason why *suf* 'drowsy' was part of a lot of repelled collocations in Table 4.6: the expected co-occurrence on the basis of their individual frequencies (which are high for both *suf* and the frequent verbs), is much higher than their observed co-occurrence. The latter example illustrates that, instead of just focusing on either the verb or the intensifier perspective, it is crucial to consider the horizontal and the vertical axis of the cross-tabulation as two intersecting dimensions.

We may furthermore wonder how we should interpret the gaps that are arbitrarily scattered across the cross-table. These gaps indicate that, for some reason, this particular verb-intensifier combination is not found in the corpus. Nonetheless, it is not necessarily the case that these combinations are strictly impossible or ungrammatical, they are merely *unattested*. First of all, it is possible that some of the combinations that are unattested in the current data set – especially combinations of frequent verbs and flexible intensifiers – do show up in other corpora. See the following examples that were found in online newspapers and magazines and do not sound strange at all.

(143) Het Surinaamse deel van het publiek **lacht zich wezenloos** om haar pittig gekruide wijsheden. (Volkskrant.nl, 2000)
[...] laughs itself blank [...]
'The Surinam part of the audience is laughing their heads off at her spicy profundities.'

(144) De jaren 80, de één blikt met weemoed terug, de ander **schaamt zich te pletter** voor die maffe kuif die toen in de mode was. (Libelle.nl, 2017)
[...] the other embarrasses himself to smithereens [...]
'The 80s, while one looks back at them nostalgically, the other is extremely embarrassed about that crazy flick that was fashionable at the time.'

Other specific combinations may be dispreferred because there already exists an alternative, very frequent, conventional combination. This brings us to the notion of statistical preemption, which assumes that a specific potential utterance may be blocked if a competing alternative utterance has been observed with some frequency in the same context. Statistical preemption has been suggested to play an important role in language acquisition, in that it helps learners to avoid syntactic overgeneralisations (see, e.g., Goldberg 1993, Braine & Brooks 1995, Brooks & Tomasello 1999). Boyd & Goldberg (2011) and Perek & Goldberg (2017) have further shown that there is also evidence of preemptive processes in productive uses of a construction in adult language use. Boyd & Goldberg (2011) find that speakers avoid using both novel and familiar *a*-adjectives in attributive position (e.g. **the asleep/ablim cow*) if they are presented with consistent use in a preemptive context (e.g. *the cow that is asleep/ablim*). The experiment in Perek & Goldberg (2017) shows that adult language users readily learn the function of a construction and are able to generalise that function to new verbs beyond the input, unless a verb is consistently observed in one construction and not another. In that case, the verb-specific

behaviour takes precedence over the function of the construction, preventing the verb's extension to an alternative construction, regardless of the functions associated with the patterns in question. Often, statistical preemption is said to operate on individual items; this also appears to be the case in the intensifying fake reflexive resultative construction. For example, if native speakers of Dutch want to intensify the verb *piekeren* 'to worry', they know from experience that the ideal candidate is *suf* 'drowsy' and that the intensifier *een hoedje* 'a little hat' should be used in combination with the verb *schrikken* 'to be startled'. That is, language users may be so familiar with the conventional combinations of *zich suf piekeren* 'to worry oneself drowsy' and *zich een hoedje schrikken* 'to startle oneself a little hat' that they are hesitant to use *piekeren* with *een hoedje*. Still, there is no grammatical rule that says that *zich een hoedje piekeren* 'to worry oneself a little hat' is ill-formed and at least some speakers would probably judge it to be more or less acceptable – irrespective of whether they would use it themselves. As a way of calculating the probability that statistical preemption will take place, Goldberg (2011) has suggested the following: in order to find out whether the use of a particular verb in one construction (CxA) is likely to be preempted by its use in a functionally equivalent, competing construction (CxB), we could use the formula P(CxB|contexts in which CxA is at least as appropriate).[40] The present study is not concerned with competing constructions, but we are interested in whether the frequent co-occurrence of two items within the same construction could preempt other combinations of the individual items within that construction. While we could apply the formula to examples of the intensifying fake reflexive resultative construction as well (e.g. P(*suf*|all uses of *piekeren* with any intensifier) is 0.09), this section clearly demonstrated that there are interactions between the verb and the intensifier slots which complicate things. Because of these interaction effects, the covarying collexeme analysis and the cross-tabulation approach were found to be more viable methods for capturing co-occurrence preferences or dispreferences in the intensifying fake reflexive resultative construction.

In addition, if we look at all the cells with frequency 1 or 2, it becomes clear that the availability of a highly frequent, conventional expression does not necessarily fully block the extensibility of the verbs or intensifiers involved. For example, both the verb *piekeren* 'to worry' and the intensifier *blauw* 'blue' are part of conventional collocations, viz. *zich suf piekeren* 'to worry oneself drowsy' and *zich blauw betalen/ergeren* 'to pay/annoy oneself blue', but we still find example (145) in the data set.

[40] Goldberg (2011: 134-136) admits that there is often no full functional overlap between two distinct constructions, and that it is important to delimit in which contexts both constructions are truly interchangeable.

(145) Een katholiek die anders dan zijn geloofbroeders uit Genève en Dordt nu eenmaal niet gewend was **zich blauw** te **piekeren** over holle doop of volle doop. (SoNaR-NL)
[...] himself blue to worry [...]
'A catholic who, unlike his brothers of the faith from Geneva and Dordt, was not used to worrying a lot about "hollow" or "full" baptisms.'

It appears that language users may sometimes (unintentionally or not) come up with "unconventional" (in the sense of infrequent) expressions by combinining a conventional (i.e. token frequent) verb with a conventional intensifier. In the case of *zich blauw piekeren* 'to worry oneself blue' the result sounds rather odd because both the verb and intensifier display a rather limited combinatorial flexibility. This is not the case for, e.g., *zich suf schrikken* 'to startle oneself drowsy', another combination which occurs only once and is therefore not strictly speaking conventional. The difference between *zich blauw piekeren* 'to worry oneself blue' and *zich suf schrikken* 'to startle oneself drowsy' is that, in the latter case, the verb and intensifier are not just highly token frequent, they are also relatively type frequent. Even though both *suf* 'drowsy' and *schrikken* 'to be startled' are part of certain conventional collocations, they also have a wide range of uses in the construction outside of these collocations. Much like the (unattested) examples of *zich wezenloos lachen* 'to laugh oneself blank/vacant' or *zich te pletter schamen* 'to embarrass oneself to smithereens' (cf. supra), their intercombination – although highly infrequent – does not really sound unfamiliar. Moreover, language users are notorious for their linguistic creativity and they have been found to intentionally flout convention. In the words of Hanks (2004) language users can "exploit" an established/conventional collocational norm (see also Chapter 2) and Tomasello (1998: 433) has suggested that "much of the creativity of language comes from fitting specific words into linguistic constructions [or in this case collocational patterns, EG] that are non-prototypical for them". The construction provides the language user with the opportunity to show off his/her linguistic creativity and distinguish himself/herself from others. We occasionally come across specific sentences in which the intensifier appears to be carefully selected because it provides a better fit with the context or because it creates a certain humorous or other special effect. In (146), for example, the intensifier *het vuur uit de sportschoenen* 'the fire out of the trainers', a variation of the more frequent intensifier *het vuur uit de sloffen* 'the fire out of the slippers', is a clever find given the sports context in which it is used.

(146) Arme scouts. **Lopen zich het vuur uit de sportschoenen** voor hun club, en dan zien ze dat er sprake is van erosie van clubliefde. (SoNaR-NL)
poor scouts. run themselves the fire out of the trainers [...]
'Poor scouts. They run their socks off for their club, only to find out that the club is not as loved as it used to be.'

More examples of this kind of linguistic creativity will be given in §4.2.2 below, as they are slightly more frequent in the Belgian data set.

Finally, it is not always easy to explain why some verb-intensifier combinations are judged to be "more acceptable" than others in the first place. In constructional approaches, the notion of semantic compatibility is often invoked to explain the distributional properties of a construction: a lexical item is more likely to be selected as a potential slot filler if it is compatible with the semantics of the construction (see, e.g., Goldberg 1995, Boas 2003 on semantic compatibility in the resultative construction). If we apply this to the interaction of slots within the same construction, we could assume that the co-occurrence of specific lexical fillers is also motivated by a semantic link between the two items. This is what Stefanowitsch & Gries (2005: 2) propose as one of the basic assumptions behind the covarying collexeme analysis: "there are constraints holding between different slots of a construction (i.e. words in such slots may covary systematically) and [...] these constraints are based on semantic coherence". They show that that is indeed the case for the verb and the resultative gerund in the *into*-causative construction. The predominant relationship in the *into*-causative is one in which the first verb expresses some kind of trickery and the resultative gerund a type of belief (e.g. *fool into thinking*, *mislead into believing*, *delude into believing,* etc.), which reflects general knowledge about the way in which our mental states can be influenced. The importance of semantic compatibility was also attested for a construction much more closely related to our current construction, viz. the Body-Part-*Off* construction (cf. §2.2.1). Sawada (2000: 364) talks about the "conceptual association" between the verbal activity and the body part and Cappelle (2014: 270) refers to "encyclopaedic relatedness" of the verbal meaning and the body part to explain why some combinations have become stronger collocations than others (see Ch2, §2.2.1). Interestingly, Cappelle also observes that in the case of the Dutch reflexive ditransitive (i.e. the intensifying fake reflexive resultative construction with NP intensifiers) "the rule seems to be that there is no obvious relationship whatsoever between the activity and the result [intensifier, EG]" (2014: 274). Of course, it needs to be pointed out that this observation was based on the verbal collocates of a small set of nominal intensifiers. If we take into account all the intensifiers in our current data set, we do find that some verb-intensifier combinations are unmistakably motivated by a semantic or conceptual link. As semantic coherence will be an important factor in determining the productivity of the verb slot in subschemas with a lexically specified intensifier, we defer the detailed discussion to the next section. On a final curious note, even in the absence of an obvious link between the intensifier and the verb, language users appear to look for some kind of explanation or motivation for these sometimes puzzling verb-intensifier conventions. When someone asked their fellow forum users where the expression "zich een hoedje schrikken" came from (to be sure, we may wonder what it is about *een hoedje* 'a little hat' that nearly restricts its collocational range to the

verb *schrikken* 'to be startled'), we find regular language users coming up with answers like the ones in (147):[41]

(147) Zo schrikken dat je je hoedje verliest.[42]
'Being so startled that you lose your little hat.'

In stripverhalen schiet dat hoedje altijd de lucht in.
'In comic books the little hat always jumps up in the air (when people are startled).'

Van schrik gaan je haren recht overeind staan; en dit lijkt op een hoedje!
'If you are scared, your hair stands up and that looks like a little hat!'

Thus, perhaps the real question is not so much whether there is an objective, encyclopaedic link between the verb and the intensifier, but whether language users can somehow conceive of a combination as semantically or conceptually motivated.

4.2.2 Synchronic variation

As a first exploration of the potential national differences in collocational preferences, we applied the hierarchical cluster analysis to the Belgian Dutch data, on the basis of the same top twenty covarying collexemes as in §4.2.1.

[41] https://www.startpagina.nl/v/kunst-cultuur/etymologie/vraag/21964/uitdrukking-schrik-hoedje (last accessed 7 November 2017).

[42] This explanation in particular is interesting from a linguistic point of view because it is at odds with the semantics traditionally associated with the formal pattern. Strictly formally speaking, *zich een hoedje schrikken* is an instance of the reflexive ditransitive construction (though see Chapter 2, §2.2.2, on why we have opted to not make this formal distinction in the present investigation) and therefore should be (literally) translated as 'to be so startled that you gain (not lose) a little hat'.

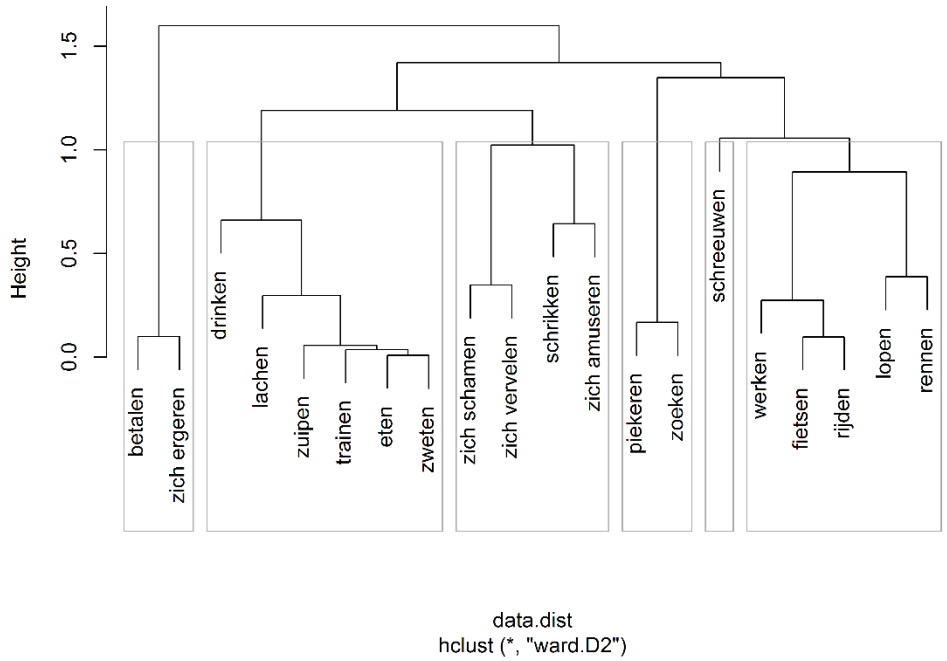


Figure 4.7. Dendrogram HCA of the top 20 verbs in SoNaR (clustered according to covarying intensifiers)

If we compare Figure 4.7 to Figure 4.5, there are a couple of similar clusters, but the overlap is only partial. While some inherently reflexive verbs are grouped together in the third cluster from the left, *zich ergeren* 'to be annoyed' is in a different cluster. A number of other verbs are found in the same clusters in both national varieties, i.e. *drinken* 'to drink' and *zuipen* 'to booze' in the second cluster from the left, *zoeken* 'to search' and *piekeren* 'to worry' in the fourth cluster from the left, and most verbs of physical activity in the rightmost cluster. Despite the substantial verbal overlap between Belgian and Netherlandic Dutch, especially among the most frequent verbs, it appears that some verbs display different collocational behaviour in the national varieties of Dutch. However, what the national varieties do share is that the majority of the clusters are not perfectly semantically motivated. While some clusters contain verbs that we would expect to group together based on their semantics, they are either incomplete or contain at least one verb that does not intuitively fit in with the others.

Figure 4.8. Dendrogram HCA of the top 20 intensifiers in SoNaR-BE (clustered according to covarying verbs)

A comparison of Figure 4.8 and Figure 4.6 suggests that some of the top intensifiers have rather similar collocational patterns in both national varieties. The fourth cluster from the left in Belgian Dutch is exactly the same as the fourth cluster in Netherlandic Dutch and there are some shared intensifiers in other clusters as well, i.e. *een ongeluk* 'an accident', *een hoedje* 'a little hat' and *wezenloos* 'blank/vacant' in the first cluster from the left, the colour adjectives in the second cluster from the left and *de pleuris* 'the pleurisy', *in het zweet* 'in the sweat' and *uit de naad* 'out of the seam' in the rightmost cluster. At the same time, there are some clear differences in the internal constitution of the clusters as well and it appears that *suf* 'drowsy' is not singled out in Belgian Dutch. If we approach these clusters from a semantic point of view, we once more find some semantic inconsistency or odd ducks in most clusters. For example, the third cluster from the left is internally homogeneous in itself, but it is missing some other prototypical negatively connoted states that are scattered across the other clusters (e.g. *wezenloos* 'blank/vacant', *dood* 'dead', *kapot* 'broken'...). Like in Netherlandic Dutch, the cluster of inalienable possession intensifiers is fully semantically motivated and complete. We will see in §4.3 that these NP+PP intensifiers retain much of their original semantics, which has an impact on the range of verbs they can co-occur with.

The exploratory comparison of the cluster analyses suggests that there are some shared collocations in both national varieties, but that there are also some national differences with respect to collocational preferences. For a further investigation of the collocational patterns in Belgian and Netherlandic Dutch, we applied the covarying

collexeme analysis to the Belgian data as well. Given the difference in size between the two data sets, we cannot compare the actual values for collostruction strength, but we can compare the ranking of the collocations in Table 4.8.

Table 4.8. Top 20 attracted and top 10 repelled covarying collexemes in SoNaR-BE

VERB	INTENSIFIER	FREQ. V	FREQ. INT	OBS. V-INT	EXP. V-INT	$\Delta P_{VERB-TO-INT}$	$\Delta P_{INT-TO-VERB}$	COLL. STR.
betalen	blauw	167	307	163	20.97	0.91	0.53	160.39
schrikken	een hoedje	465	200	197	38.04	0.42	0.87	153.82
werken	uit de naad	346	192	140	27.17	0.38	0.64	86.03
zich amuseren	rot	181	241	118	17.84	0.6	0.46	84.56
zich ergeren	blauw	232	307	138	29.13	0.52	0.41	75.04
piekeren	suf	65	129	57	3.43	0.85	0.44	69.34
zich schamen	dood	43	177	31	3.11	0.66	0.17	26.61
zich ergeren	dood	232	177	65	16.8	0.23	0.29	24.66
schrikken	een aap	465	32	31	6.09	0.07	0.79	21.28
lopen	het vuur uit de sloffen	125	24	19	1.23	0.15	0.75	20.6
lopen	de benen PREP het lijf	125	25	19	1.28	0.15	0.72	19.99
drinken	een stuk in de kraag	31	13	11	0.16	0.35	0.84	19.85
schreeuwen	schor	29	14	11	0.17	0.38	0.78	19.57
zich vervelen	steendood	78	16	14	0.51	0.18	0.85	19.41
lachen	een breuk	123	26	18	1.31	0.14	0.65	17.86
schrikken	een bult	465	53	39	10.08	0.08	0.56	17.47
schrikken	een ongeluk	465	57	40	10.84	0.08	0.52	16.7
zoeken	suf	44	129	21	2.32	0.43	0.15	15.67
lachen	een kriek	123	11	11	0.55	0.09	0.95	14.47
werken	in het zweet	346	21	19	2.97	0.05	0.77	14.12
werken	blauw	346	307	2	43.44	-0.14	-0.15	18.6
schrikken	te pletter	465	445	25	84.63	-0.16	-0.16	18.07
werken	rot	346	241	1	34.1	-0.11	-0.15	15.22
zich ergeren	te pletter	232	445	8	42.22	-0.16	-0.09	11.88
zich ergeren	rot	232	241	4	22.87	-0.09	-0.09	6.54

werken	dood	346	177	7	25.05	-0.06	-0.11	5.37
schrikken	kapot	465	72	3	13.69	-0.03	-0.15	3.69
lachen	een hoedje	123	200	1	10.06	-0.08	-0.05	3.58
zich vervelen	blauw	78	307	1	9.79	-0.12	-0.03	3.52
lopen	dood	125	177	1	9.05	-0.07	-0.05	3.14

There are ten verb-intensifier combinations that are featured in the top twenty of attracted collexemes in both national varieties of Dutch (indicated in grey in the table), although there is some variation in their ranking. The national variation in intensifier preferences that was discussed in §4.1.2.2 does appear to explain some of the variation in these top collexemes, but not all of it. Regardless of national preferences, most of the intensifiers in the top collocates are relatively token frequent in both Belgian and Netherlandic Dutch. Some of the non-overlapping strong collocations obviously feature intensifiers that were found to be (nearly) exclusive to one national variety, e.g. *zich groen en geel ergeren* 'to annoy oneself green and yellow' and *zich wild ergeren* 'to annoy oneself wild' in Netherlandic Dutch and *zich een aap schrikken* 'to startle oneself a monkey' and *zich steendood vervelen* 'to bore oneself stone-dead' in Belgian Dutch. Remarkably, there are no shared repelled collexemes (i.e. the bottom ten collexemes) between Belgian and Netherlandic Dutch and there are none with *suf* 'drowsy' in Belgian Dutch either. A closer look at the repelled collexemes in Belgian Dutch once more shows that the covarying collexeme analysis fails to grasp some of the intricacies of this particular construction. For example, *zich een hoedje lachen* 'to laugh oneself a little hat' is signalled as one of the top repelled collexemes. This means that, based on the relatively high token frequency of *een hoedje* 'a little hat' and *lachen* 'to laugh' individually, we might have expected the two elements to co-occur more frequently than they did. However, if we take into account that *een hoedje* 'a little hat' has a type frequency of only 4 verbs in total, and that one of these verbs, viz. *schrikken* 'to be startled' accounts for 197 of the 200 tokens, we could actually posit that it is surprising that it appears with another verb aside from *schrikken* 'to be startled' at all (cf. the phenomenon of statistical preemption supra, but see §4.3.2 for some nuance). In any case, a cross-tabulation approach which includes information on type frequency, hapax count, type-token distribution and relative entropy more adequately captures these kind of subtleties in the data. See Table 4.9 for the cross-tabulation of the top 15 intensifiers and top 15 verbs in the Belgian Dutch data.

Table 4.9. Full cross-tabulation for the top 15 verbs/intensifiers in SoNaR-BE

	te pletter	blauw	rot	een hoedje	uit de naad	dood	suf	de ziel uit het lijf	kapot	een ongeluk	een bult	de pleuris	de longen uit het lijf	een aap	krom	Tokens	Types	Hapaxes	H_{rel}
schrikken	25		82	197		21			3	40	39	3		31		465	20	7	0.414
werken	96	2	1		140	7		11	17			10		1	16	346	24	11	0.414
zich ergeren	8	138	4			65			4							232	15	8	0.264
zich amuseren	54		118						6							181	5	1	0.182
betalen		163													1	167	4	2	0.030
lopen	15				16	1	2	15	1			6	12			125	24	12	0.555
lachen	34		7	1		10			6	2	9	1			7	123	20	9	0.526
zich vervelen	27	1	8			22			3							78	9	4	0.353
rijden	14				17	1	1	17	1			6			1	69	17	12	0.460
piekeren			2				57		1							65	7	4	0.128
zoeken	13		5				21					2			1	44	7	3	0.298
zich schamen	3		2			31			6							43	5	1	0.203
drinken	7								2							31	9	3	0.403
schreeuwen	1						1	3	1				11			29	7	4	0.314
zweten	20					2		1	4							28	5	2	0.207
Tokens	445	307	241	200	192	177	129	98	72	57	53	39	38	32	27				
Types	74	7	19	4	14	22	37	27	28	16	3	16	11	2	6				
Hapaxes	42	4	10	3	9	13	26	15	17	13	0	10	5	1	4				
H_{rel}	0.623	0.157	0.282	0.18	0.212	0.399	0.463	0.533	0.558	0.275	0.146	0.462	0.374	0.027	0.224				

Again, we can observe that a large proportion of the uses of the intensifying fake reflexive resultative construction (in newspaper data) are conventional: the combinations of the top 15 verbs and intensifiers cover 72% of the entire data set (which is 7% more than in Netherlandic Dutch). Based on the randomly scattered empty and filled spots across the board, there is a lot of variation in the combinatorial flexibility of the top verbs and intensifiers in both national varieties. On the horizontal axis, we find a lot of similarities with Netherlandic Dutch. For one, *schrikken* 'to be startled' is the most frequent verb overall, but *werken* 'to work' is the most flexible verb with 24 different intensifier types. In Belgian Dutch, *werken* has to share its top spot with *lopen* 'to run', which also occurs with 24 types – half of which are hapaxes – and has the highest relative entropy of all verbs, hinting at a relatively balanced distribution. Verbs like *betalen* 'to pay' and *piekeren* 'to worry' have an extremely skewed distribution and a much more limited collocational range in both national varieties: their profiles are characterised by one high-frequency cell and a lot of gaps. If we look at the collocational behaviour of the top 15 intensifiers, we also observe remarkable national consistency in the patterning. On the one hand, there are a number of intensifiers that are both relatively token frequent and distributionally flexible, e.g. *suf* 'drowsy', *rot* 'rotten', *kapot* 'broken', *dood* 'dead'; on the other, we find intensifiers that are only token frequent by virtue of being part of a frequent conventional collocation, e.g. *een hoedje* 'a little hat', *blauw* 'blue', *het vuur uit de sloffen* 'the fire out of the slippers'. Perhaps the biggest difference is that in Belgian Dutch, there is one intensifier that clearly outperforms its "competitors" on all fronts, viz. *te pletter* 'to smithereens'. In Netherlandic Dutch, only *suf* 'drowsy' behaves similarly in terms of type frequency and hapax count, but it is not the most token frequent intensifier in the data set. Interestingly, the intensifiers with a significant preference for one of the two national varieties (cf. supra) are not necessarily also more flexible in that variety: *kapot* 'broken' is relatively speaking much less frequent in Belgian Dutch, but it shows remarkable flexibility in terms of type frequency and hapax count. The information in this cross-table also provides us with a possible explanation for why *een hoedje* 'a little hat' can pair up with *lachen* 'to laugh', despite its near-exclusive association with *schrikken* 'to be startled'. The choice for *lachen* 'to laugh' in particular, could be motivated by its type frequency and type-token distribution: *lachen* is found with 16 other intensifiers and has one of the higher relative entropies, so language users may feel like they can use this verb with pretty much any intensifier, even *een hoedje* 'a little hat'. Finally, it is noteworthy to point out that some of the intensifiers that were almost exclusively found in one of the two national varieties show very little flexibility. For Netherlandic Dutch, for instance, we see that *groen en geel* 'green and yellow' only combines with *zich ergeren* 'to be annoyed' and that *wild* 'wild' is split evenly between *schrikken* 'to be startled' and *zich ergeren* 'to be annoyed'; in Belgian Dutch *steendood* 'stone-dead' is only found to co-occur with *zich vervelen* 'to be bored' and *een aap* 'a monkey' is nearly exclusively associated with *schrikken*

'to be startled'. This suggests that it is perhaps not so much the intensifier that is nationally exclusive, but the entire collocation: it is likely that Belgian and Netherlandic Dutch have gone through different conventionalisation processes.

If we gloss over the behavioural details of specific verbs and intensifiers and consider the overall intersections of the horizontal and vertical axes, we must conclude that Belgian and Netherlandic Dutch share a lot of conventional verb-intensifier combinations, e.g. *zich uit de naad werken* 'to work oneself out of the seam', *zich blauw betalen* 'to pay oneself blue', *zich een hoedje schrikken* 'to startle oneself a little hat', *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers', *zich suf piekeren* 'to worry oneself drowsy', etc. Of course, just like speakers of Netherlandic Dutch, Belgian speakers may also choose to ignore the conventional combinations and exploit the productivity of the construction to come up with creative alternatives. In (148), the specific verb-intensifier combination really paints a picture of a woman sitting in her kitchen all day, bored out of her mind. This effect would be largely lost if the journalist had used a more run-of-the-mill intensifier like *dood* 'dead' instead.

(148) Vader heeft een belangrijke job in een grootwarenhuisketen en een liefje in Oxford, moeder **verveelt zich de poten vanonder haar keukenstoel**. (SoNaR-BE)
[...] mother bores herself the legs from under her kitchen chair
'Father has an important job at a department store and a girlfriend in Oxford, and meanwhile mother is bored out of her mind (at home, in her kitchen).'

Example (149) works in a similar way. The example conjures a vivid image of little bugs suddenly scuttering about when they are disturbed, which would not be the case if the journalist had opted for, e.g., the conventional intensifier *een hoedje* 'a little hat'.

(149) De gierigaard leeft als een pissebed: verscholen onder een stoeptegel . Wanneer iemand de steen optilt en de gierigaard aanwijst, **schrikt** die **zich een pissebed.** (SoNaR-BE)
[...] startles that one himself a sow bug
'The miser lives like a sow bug: hidden under a paving stone. When someone suddenly lifts it and points at him, he is very startled.'

Example (150) contains a word play on the conventional combination of *zich blauw betalen* 'to pay oneself blue' and the colour associated with the Belgian political party, the VLD (the Flemish Liberals and Democrats). These kinds of clever puns obviously only work if the audience has the necessary background information.

(150) Over de fiscale supertarieven die voor vakantiegeld, eindejaars- en andere premies gelden, zullen we het maar niet hebben. We **betalen ons** dus nog altijd **VLD-blauw**. (SoNaR-BE)
[...] we pay ourselves thus still VLD-blue
'Let's not talk about the fiscal super-rates that apply to vacation pay, annual or other bonuses. We are still paying a lot of money.'

The examples above, especially (148) and (149), illustrate the phenomenon of constructional coercion: regardless of the specific lexemes that are filled into the slot, the language user can rely on his linguistic experience and the constructional semantics to make sense of the sentence. In the next sections, we will investigate how the (usually subtle but sometimes more substantial) national differences that were outlined in this section influence the degree of productivity of the different subschemas in the network and, importantly, whether these differences are significant enough to warrant separate network representations for the two national varieties of Dutch.

4.3 Productivity in the constructional network

In Chapter 2, we discussed the theoretical notion of constructional productivity, which was defined as the extensibility of the schema, i.e. the way in which specific slots in a construction can be extended to new types. In previous sections, we already touched upon several factors that may play a role in the degree of productivity of a construction, e.g. type and token frequency, hapax count, skewness of the distribution (H_{rel}) and collocational constraints. In this section, we will elaborate on the empirical measures of productivity that were introduced in Chapter 2 and apply them to the intensifying fake reflexive resultative construction. In what follows, we will argue that productivity should be measured at different levels of the hierarchy in the constructional network. Productivity is tightly intertwined with schematicity in that a more productive schema, which is subject to less constraints, is more schematic (or abstract), and, accordingly, situated at a higher level in the hierarchy (Ch2, §2.1.2 and §2.1.3).

Based on the results of the investigation so far, it would appear that the construction at the highest level of schematicity [SUBJ V REFL INT] is very productive. Even though the verb slot is somewhat constrained in as far as the verb must involve some experience or activity that is eligible for intensification (cf. §4.1 supra), there is still an enormous variety of different verbs that can be used in the construction. The intensifier slot also displays a wide range of different intensifiers: although the slot shows a clear *preference* for items that have a negative connotation in their original lexical semantics, we found no immediate restrictions pertaining to the intensifier slot at this level. That is not to say that any Dutch word can be filled in the intensifier slot, but no clear constraints can be formulated. However, if we focus our attention on the lower level of subschemas in which either the verb or the intensifier are lexically specified, we get a very different picture. The collexeme analysis and cross-table in the previous section already showed that there are clear differences in the combinatorial flexibility of certain verbs and intensifiers in Dutch. For example, the verb *werken* 'to work' was found with a lot of different intensifier

types, many of which were hapaxes. In this section and the next, we will show that we can interpret this as *werken* 'to work' forming a partially specified subschema with a productive intensifier slot, viz. [SUBJ *werken* REFL INT], in the sense that (i) it hosts a lot of types already, (ii) it is likely to attract new types, and (iii) it is not subject to any obvious semantic constraints. The verb *piekeren* 'to worry', on the other hand, primarily co-occurs with the intensifier *suf* 'drowsy'. We will see in the next section that, in the constructional network, this combination takes the shape of a lexically specified micro-construction [SUBJ *piekeren* REFL *suf*]. Given the limited combinatorial flexibity of *piekeren*, it is unlikely that we can posit a productive subschema [SUBJ *piekeren* REFL INT] but, given the variety of verbs that appear with *suf* 'drowsy', a productive subschema [SUBJ V REFL *suf*], in which the intensifier – rather than the verb – is lexically specified, is much more likely. Other intensifiers display varying degrees of combinatorial flexibility. The intensifier *een slag in de rondte* 'a punch around' was found to co-occur with a variety of activity verbs in Table 4.7, but the experience verbs are curiously absent from its collocational range (cf. infra). This collocational preference is likely to have an impact on the productivity of the subschema [SUBJ V REFL *een slag in de rondte*], which may require an extra specification on its verb slot. In addition, given the (almost) exclusive association of *een hoedje* 'a little hat' with *schrikken* 'to be startled' or *groen en geel* 'green and yellow' with *zich ergeren* 'to be annoyed', it is doubtful whether the network contains productive intensifier-specific subschemas [SUBJ V REFL *een hoedje*] or [SUBJ V REFL *groen en geel*] (but it does contain the verb-specific subschemas [SUBJ *schrikken* REFL INT] and [SUBJ *ergeren* REFL INT]). In the following two paragraphs, we will argue that it is crucial to look at productivity at intermediary levels of abstraction, because productivity at the highest level of schematicity does not necessarily entail that all lower-level subschemas are also (equally) productive. In fact, the complex structure of the constructional network, with both open, productive subschemas and lexically specified micro-constructions, reflects how the construction presents itself as both productive and constrained by convention at the same time, as we have repeatedly stated throughout this thesis.

In this study, we propose a multidimensional model of productivity which incorporates both the frequency-based measures of Baayen and colleagues and the constructional productivity model of Barðdal. In Chapter 2, we already briefly introduced the productivity model of Baayen, which is based on the idea that productivity is a complex phenomenon with different facets that need to be measured and evaluated separately. Baayen & Lieber (1991) and Baayen (2009: 901) state that type frequency (V) highlights the "past achievement" of a word formation pattern (in our case, a construction) and gives an indication of the current extent of use, i.e. the REALISED PRODUCTIVITY, rather than of the extensibility of the pattern (see also Bauer 2001: 48, who

similarly argues that "type frequency is the result of past productivity rather than an indication of present productivity").[43]

Realised productivity = V	V = types

In order to estimate the extensibility of a construction, we need to take into account the so-called hapax legomena (HL), i.e. the types that occur only once in the corpus. By taking the ratio of HL of a specific slot of a construction, e.g. the verb slot, to the total number of tokens of that specific slot in the construction, we get a measure for the POTENTIAL PRODUCTIVITY $\mathscr{P}$ of that constructional slot. This gives an indication of the probability that we will encounter new types in the pattern as the total sample of tokens is increased. Although Baayen does not mention this himself, the increase of sample size could perhaps be interpreted in terms of linear time: $\mathscr{P}$ indicates the probability that we will encounter new types in the slot in the (immediate or very near) future (but see Ch5, §5.3.1 and Ch6, §6.2.2 for discussion).

Potential productivity $\mathscr{P}$ = n1/N	n1 = hapaxes; N = tokens

As Baayen & Lieber (1991: 810) point out, potential productivity is positively influenced by a high number of hapaxes (i.e. a high denominator) and a low number of high-frequency tokens that may inflate the nominator. However, in contrast to what has generally been assumed, the existence of highly frequent tokens does not necessarily constrain the productivity of the relevant construction: a pattern may have a number of frequent "prototypical" types, while at the same time producing a large number of one-offs as well. In fact, this is exactly what Baayen and Lieber (1991: 832-836) find for the English prefix *re-*, which shows a combination of a small number of high-frequency formally and/or semantically idiosyncratic words (e.g. *remove, recover, recall, react…*) and a large number of fully compositional hapaxes (e.g. *reheat, reforest, repoint, retake…*). For that reason, we would like to integrate an additional measure of productivity into our multidimensional model. The measure is very similar to potential productivity but less

[43] Baayen and Lieber (1991: 818) argue that the number of observed types is determined by other factors as well, and it is not always clear how these factors interact with productivity. For one, the pragmatic usefulness may explain why one word-formation rule produces more types than another. The suffix *-erd* in Dutch combines with adjectives to coin personal names with a fairly negative connotation, e.g. *mallerd* 'silly one', *stommerd* 'fool', *natterd* 'wet one'. In theory, the suffix is productive in that it could be attached to any adjective, but the contexts in which one would use a word like that are limited (see also Baayen 1990). This also relates to the semantic flexibility of the word-formation process; compounding, for instance, is such a semantically versatile and neutral formation pattern that it can produce an almost infinite number of new words.

sensitive to high-frequency items[44]; instead of taking the proportion of hapaxes to the total number of tokens, it may be illuminating to consider how many of the total number of *types* are represented by hapaxes (which is how Bardal (2008: 26) reinterprets the global productivity measure by Baayen, cf. infra).

Hapax/type ratio = n1/V	n1 = hapaxes; V = types

It is important to observe that the absolute value of these ratios does not give information on *the* degree of productivity of a construction. Potential productivity (and, by extension, the HAPAX-TYPE RATIO) is intended as a measure for comparing the relative degrees of productivity of two or more patterns but there are no fixed cut-off points to discriminate "low" from "high" productivity. Concretely, given data sets of a more or less similar size (cf. infra), if the ratio of hapaxes of pattern A to the total frequency of pattern A is higher than the ratio of hapaxes of pattern B to the total token frequency of pattern B, pattern A is more productive than pattern B.

Now that we have illuminated how type frequency (V) and potential productivity ($\mathscr{P}$) highlight different aspects of the productivity of a (constructional) pattern, the next step is to bring them together in a model that allows us to say something about the global productivity of that pattern. Baayen and Lieber (1991: 818) propose a bi-dimensional visualisation of GLOBAL PRODUCTIVITY, which they describe as follows:

> Global productivity ($\mathscr{P}$,V)
> = the global productivity P* of a word-formation rule [or construction, EG] can be summarised in terms of its coordinates in the $\mathscr{P}$-V plane, with the degree of productivity on the horizontal axis and the extent of use V on the vertical axis.

By plotting multiple patterns in this $\mathscr{P}$-V plane, it is possible to quickly gauge the differences in productivity between different (sub)schemas: the (globally) more productive (sub)schema is situated more to the top right with both high $\mathscr{P}$- and V-values while the less productive (sub)schema, which has lower $\mathscr{P}$- and V-values, will be situated to the bottom left of the plane. However, there are some limitations to this concept of P* as well, in that it remains difficult to assess which aspect should be given more weight when the differences along the two dimensions do not point in the same direction. If, for instance, pattern A has a higher realised productivity (V) but a lower potential productivity ($\mathscr{P}$) than pattern B, global productivity alone does not help decide which of the two is the more productive pattern "overall" (Baayen & Lieber 1991: 818; Baayen 1993:

[44] In general, the $\mathscr{P}$-measure is not of much use in very large samples. If the sample size N continues to increase, the proportion of hapaxes will become negligible and $\mathscr{P}$ will approximate zero, even for productive affixes (Baayen and Lieber 1991: 837)

190). This is however not considered to be a major issue of concern if productivity is conceptualised as a multidimensional complex, as we will show below. As was argued and illustrated in Chapter 2, the frequency-based measures of Baayen and colleagues can be fruitfully applied to constructions and suffice as a first indication of the degree of productivity of schemas and subschemas (Zeldes 2012).

It has also been suggested that the shape of the distribution, i.e. the way in which the tokens are distributed across the different types, plays a role in determining (quantitative) productivity. Productivity is said to arise from skewed distributions, characterised by a few highly frequent types and a large number of infrequent types (like hapax legomena) (Zeldes 2012: 207-208, Gries 2012: 503-504). The idea behind this is that the many infrequent types strengthen the abstract construction. Conversely, if a construction has a more balanced distribution (i.e. in which the tokens are distributed more or less equally over a limited number of types), the construction may become exclusively associated with those types and not (or no longer) be considered to be extensible to new types. In the previous section on collocational patterns, the measure of relative entropy was introduced to quantify the variability and the shape of the distribution: low relative entropy values are indicative of a highly skewed distribution, higher values point to a more balanced distribution. We will illustrate that, while the shape of the distribution may well be of importance in the productivity of a schema, relative entropy may not be the best measure to capture the relevant aspects of the distribution and will therefore not be included in the model.

Within a constructional approach to language, in which constructions are defined as form/meaning-pairings, we should also take into account the potential influence of semantics on the productivity of constructions (i.e. qualitative rather than quantitative productivity). Barðdal (2008) puts forward an alternative model that factors in the type frequency and the semantic coherence of a schema, which, taken together, should accurately predict a schema's productivity. The argumentation of this model was set out in Chapter 2, §2.1.3, so we only briefly repeat the main hypothesis. Barðdal (2008) argues that there is an inverse correlation between the type frequency of a construction and its degree of semantic coherence. This means that the importance of semantic coherence for the productivity of a schema increases as its type frequency decreases. This is graphically represented in Figure 4.9.

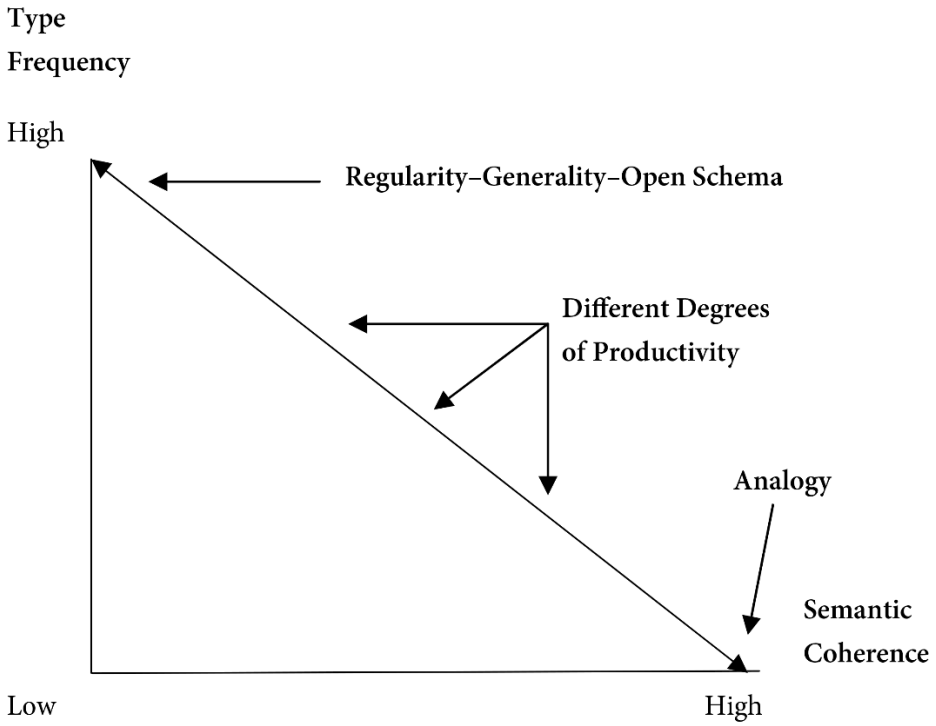


Figure 4.9. The inverse correlation between type frequency and semantic coherence (after Barðdal 2008: 38)

At the rightmost bottom of the continuum, token frequency may come in as an important factor for productivity, in contrast to what has generally been assumed (cf. supra, Bybee 1985, 1995, Bybee & Thompson 1997, Clausner & Croft 1997). In a constructionist account that accepts constructions existing at different levels of schematicity, there is no qualitative difference between analogy and productivity – the difference is simply a matter of degree (Barðdal 2008: 91). If one of the members of a schema is highly token frequent, it may come to serve as a model for analogical extensions and attract new members to the schema, eventually giving rise to (partially) productive schemas that are situated higher on the cline.

4.3.1 Synchronic use

4.3.1.1 A frequency-based productivity complex

If we first look at the productivity of the verb slot and the intensifier slot at the maximum level of schematicity, i.e. of the [SUBJ V REFL INT] pattern, we see that the verb slot is more productive than the intensifier slot, both in terms of realised productivity and in terms of potential productivity. In other words, there are more verb types (V=137) than intensifier types (V=68) and the verb slot has a higher potential for attracting new types (n1=86, $\mathscr{P}$=0.08, hapax/type=0.51) than the intensifier slot (n1=23, $\mathscr{P}$=0.02, hapax/type=0.32). It could be presumed that the creation of new intensifier types

requires more effort from the language user than extending an existing intensifier-pattern to a new verb. We have already given multiple examples of creative, one-off intensifiers serving some kind of rhetorical purpose, but we generally do not get the same effect if we extend a "conventional" intensifier to a previously unattested verb – unless the verb-intensifier combination as a whole explicitly flouts convention. At any rate, we hope to have shown so far that such broad generalisations are not very telling. Taking into consideration the observed discrepancies in combinatorial flexibility between individual verbs and intensifiers, we argue that it is much more informative to measure the productivity at lower levels in the network. It has been suggested that low-level subschemas may be more relevant for capturing the essential distributional information and the subregularities of a construction than higher-level schemas (Langacker 1999, see also Ch6, §6.2.2 and 6.2.3 for discussion). Given the interaction and covariation between verb and intensifier, it is illuminating to approach these lower-level subschemas from two angles, viz. by focusing on the productivity of the verb slot in a subschema in which the intensifier is lexically specified while at the same time considering the productivity of the intensifier slot in a subschema in which the verb is lexically specified.

(a) Intensifiers

In Table 4.10 we calculated the frequency-based measures for the top 15 intensifiers in the Netherlandic Dutch data set.

Table 4.10. Frequency-based productivity measures for the top 15 intensifiers in SoNaR-NL

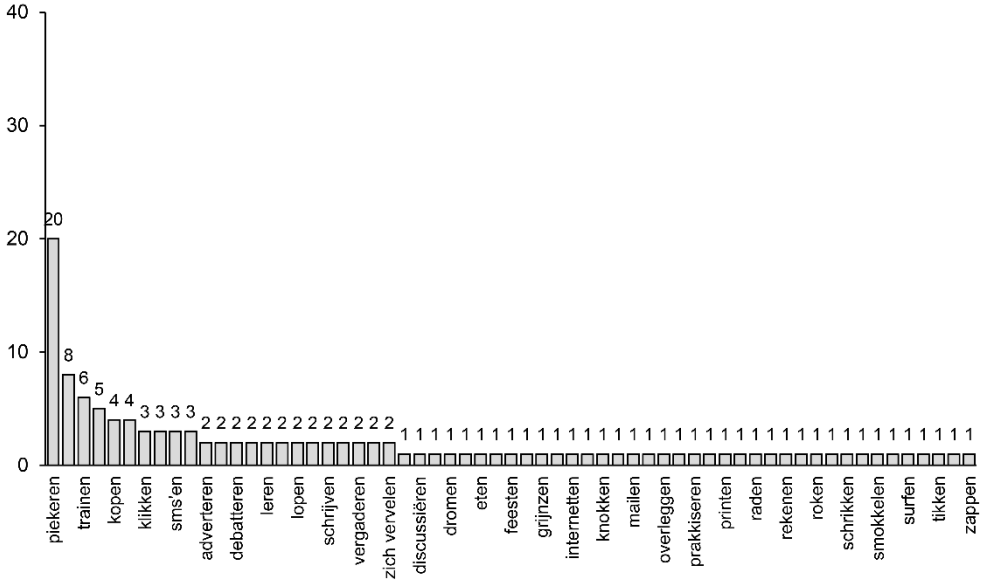
	N	V	N1	$\mathscr{P}$	HAPAX/TYPE	H_{REL}
rot	154	14	5	0.03	0.36	0.353
suf	123	61	38	0.31	0.62	0.751
dood	116	13	4	0.03	0.31	0.383
kapot	104	16	9	0.09	0.56	0.441
te pletter	60	22	18	0.30	0.82	0.47
een ongeluk	60	23	18	0.30	0.82	0.473
groen en geel	44	1	0	0.00	0.00	0
een slag in de rondte	37	22	19	0.51	0.86	0.516
wild	34	3	1	0.03	0.33	0.161
blauw	28	7	3	0.11	0.43	0.297
een hoedje	28	2	1	0.04	0.50	0.031
wezenloos	22	9	6	0.27	0.67	0.341
uit de naad	21	6	4	0.19	0.67	0.212
het vuur uit de sloffen	13	1	0	0.00	0.00	0
de longen uit het lijf	12	8	6	0.50	0.75	0.393

The previously observed exclusivity of *zich groen en geel ergeren* 'to annoy oneself green and yellow' and zich *het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers'

translates into a score of 1 for realised productivity and of zero for the other productivity measures. However, the existence of a conventional collocation does not necessarily negatively impact the productivity of an item; it appears that highly frequent tokens may influence productivity in different ways. On the one hand, we find examples for which it does appear to be the case that highly token frequent collocations are blocking – or at least heavily restricting – the extension of these intensifiers to new verbs (cf. statistical preemption, §4.2.1). In addition to the "unproductive" intensifiers *groen en geel* and *het vuur uit de sloffen*, we also find low productivity scores for *een hoedje* 'a little hat' or *wild* 'wild', which are part of the collocations *zich een hoedje schrikken* 'to startle oneself a little hat', *zich wild ergeren* 'to annoy oneself wild' and *zich wild schrikken* 'to startle oneself wild'. On the other hand, even though 14 of the 37 tokens of *een slag in de rondte* 'a punch around' are with the verb *werken* 'to work' (see Table 4.7), this does not prevent the intensifier from being used with 21 other verbs as well. We will come back to this in the section on network hierarchy below, where we will show that the existence of a frequent micro-construction like [SUBJ *werken* REFL *een slag in de rondte*] is not incompatible with a productive higher-order subschema [SUBJ V REFL *een slag in de rondte*]. The intensifier *suf* 'drowsy', has the highest degree of realised productivity with 61 verb types, more than half of which are hapax legomena, and it has one of the highest values of $\mathscr{P}$, suggesting that it has a high likeliness of occurring with even more verb types. Only *een slag in de rondte* 'a punch around' has a higher $\mathscr{P}$-score with 19 hapaxes out of 37 tokens and 22 types.

Suf 'drowsy' also has the highest relative entropy of all the intensifiers, meaning that its distribution is not primarily dominated by highly frequent verb types. Most of the intensifier-specific subschemas that score high on potential productivity also have a relatively high relative entropy value. The intensifiers with a distribution that is dominated by one or two verbs, e.g. *een hoedje* 'a little hat' or *wild* 'wild', have remarkably lower $\mathscr{P}$-scores. While this is wat we would expect on the basis of the traditional assumption that highly frequent tokens detract from productivity – although we have just shown that this is not *necessarily* the case –, it is somewhat surprising because, if we are to follow Gries (2012), skewed distributions with *lower* entropy should feed into productivity (cf. supra). In light of that, we may question to what extent the relative entropy measure adequately captures all relevant aspects of the distribution. Relative entropy conflates type frequency and the predictiveness of the distribution, and, as such, does not necessarily correlate with the other frequency-based measures that only take into account type or token frequencies or hapax counts. For one, the H_{rel} is arguably not very informative for items that have a very limited collocational range – that is, can we really speak of a "distribution" if it consists of only two or three verbs? Second, it remains somewhat unclear how skewness is really defined. The intensifier *suf*, for example, has a very high H_{rel} score because, aside from the collocation with *piekeren* 'to worry', it is indeed rather equally distributed over the other verb types. However, as those verb types

are highly infrequent and many of them are hapaxes, *suf* 'drowsy' does actually present a skewed distribution following Zeldes's (2012: 207) definition: "skewed distributions, with a roughly Zipfian frequency spectrum including very many rare types and a few frequent types", see Figure 4.10.


40
30
20
10
0
20
8
6
5
4 4
3 3 3 3
2 2 2 2 2 2 2 2 2 2 2 2 2 2
1 1 1

types. An example of this is *rot* 'rotten', which given its high token frequency of 154 has a very low $\mathscr{P}$-score of 0.03, even though about a third of the types are hapaxes. Looking back at the cross-table in Table 4.7, we see that *rot* indeed occurs particularly frequently with *schrikken* 'to be startled', which accounts for about half of all tokens, but there are also a couple of infrequent types as well. Still, while the hapax-type ratio does put the low $\mathscr{P}$-score of some intensifiers somewhat into perspective, intensifiers like *rot* 'rotten' and *dood* 'dead' do have remarkably lower type and hapax counts compared to some of the other intensifiers (e.g. *suf* 'drowsy', *te pletter* 'to smithereens', *een ongeluk* 'an accident'). Given that *rot* 'rotten' and *dood* 'dead' are intuitively omnipresent, semantically unconstrained intensifiers, this may be somewhat surprising at first, but the low productivity scores could in a way be related to their omnipresence. The high token frequencies of *rot* 'rotten' and *dood* 'dead' come from their use with a number of frequent verbs, but their low type frequencies and hapax counts indicate that they are not always the first intensifiers that come to mind when speakers want to intensify a less frequent verb. That is, the fact that they enter into these highly frequent, conventional collocations may restrict the extensibility of the intensifier to new or infrequent verbs, see the distribution of *dood* 'dead' in Figure 4.11.

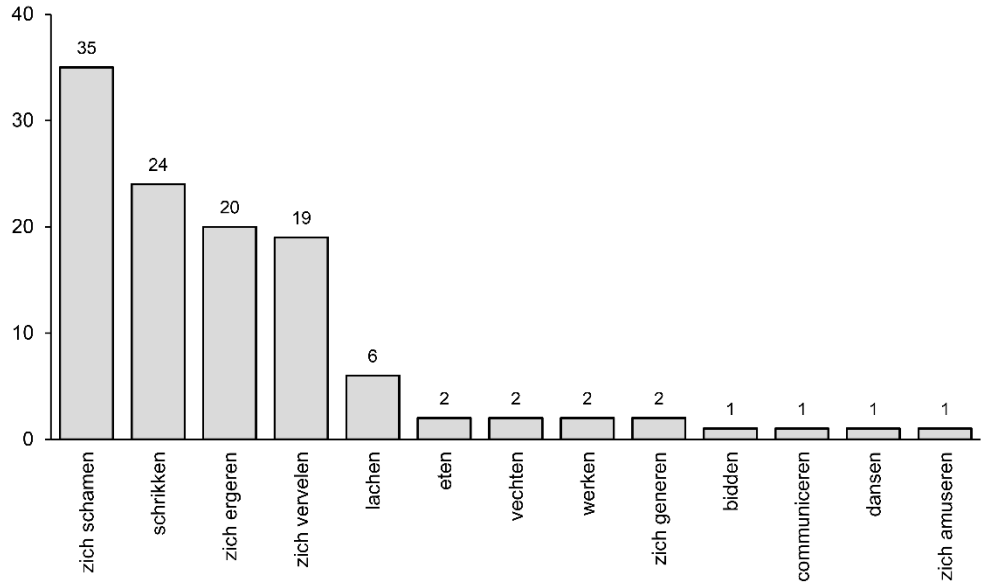


Figure 4.11. Verb distribution of *dood* in SoNaR-NL

In this case, the more or less equal distribution over a number of token frequent verb types in the left part of the graph and the relatively short tail on the right could eventually cause *dood* 'dead' to become exclusively associated with those frequent verbs (cf. supra, Zeldes 2012: 207-208). We saw the exact opposite for *suf* 'drowsy', which has low token frequencies with the most frequent verbs (cf. §4.2.1 supra) but a high type and hapax count, as was obvious from Figure 4.10. Although this again goes to show that the shape of the distribution may have an important influence on the extensibility of a construction, it casts more doubt on the validity of the relative entropy measure. If we

compare the distributions in Figure 4.10 and Figure 4.11, it is surprising that *dood* 'dead' has a lower H_{rel} value, i.e. a supposedly more skewed distribution, than *suf* 'drowsy' (cf. supra). Given the issues above, we decide not to include H_{rel} in our productivity model and we will no longer apply the measure in the remainder of this thesis.

The global productivity graph in Figure 4.12, which brings together the potential productivity on the X-axis and the realised productivity on the Y-axis, offers an instructive visualisation of the overall productivity of the intensifiers.

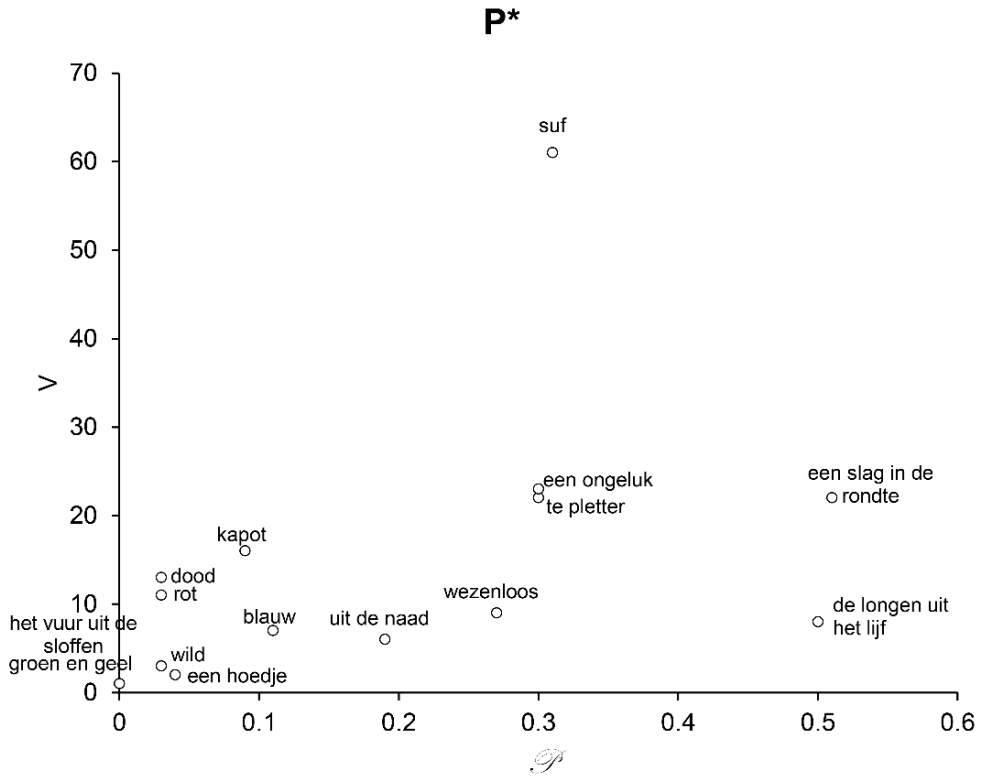


Figure 4.12. Global productivity (P*) of the top 15 intensifiers in SoNaR-NL

Based on their position on the P-V plane, we could try to group the intensifiers together in several "productivity clusters". At the leftmost bottom of the graph, we find the intensifiers that have a very low overall degree of productivity in that they do not occur with a large variety of types and are attested with a relatively low number of hapaxes, suggesting that they are not very likely to attract new verb types. In this category, we have *het vuur uit de sloffen*, *geel en groen*, *wild* and *een hoedje*. We then have a mixed category that is slightly more productive either in terms of realised productivity (i.e. they appear with a larger variety of types, see *dood*, *rot* and maybe *kapot*) or in terms of potential productivity (i.e. they are available to be extended to new types, see *blauw*, *uit de naad* and maybe *wezenloos*). In the middle of the graph, we have a small group of intensifiers, made up of *te pletter* and *een ongeluk*, which have already managed to reach some degree of realised productivity, but also score well on potential productivity. At approximately the same position on the X-axis, but situated much more to the top of the graph, we find *suf*;

at approximately the same position on the Y-axis, but situated much more to the right, we see *een slag in de rondte*. Finally, we find *de longen uit het lijf*, which has a rather low type frequency at the moment but, given the proportion of hapaxes, may well be extended to new types.

We could now try to answer the question as to which of these 15 intensifiers is the most productive *overall*. The global productivity graph is especially useful for comparing patterns for which $\mathscr{P}$ and V correlate, but it does not provide an ultimate answer if $\mathscr{P}$ and V point in opposite directions. Because of this specific drawback, recognised by Baayen (1992: 192) himself, the approach has been criticised by Gaeta & Ricca (2006: 61), among others. Within a multidimensional approach to productivity, however, it is not really an issue that items have divergent values for the different facets of productivity. If we compare *suf* 'drowsy' (highest V, lower $\mathscr{P}$) and *een slag in de rondte* 'a punch around' (lower V, highest $\mathscr{P}$) with each other, we cannot decide which of the two is most productive overall but we can of course say that *suf* has a higher degree of realised productivity than *een slag in de rondte*, whereas the latter displays a higher potential to be extended to new types.

It has been pointed out that comparing $\mathscr{P}$-values for samples that are very different in size can be problematic. If applied to a corpus or a data set in which the token frequencies of the compared items differ greatly, $\mathscr{P}$ tends to be overestimated for the lower frequency items (Gaeta & Ricca 2006). Gaeta & Ricca (2006: 63) and Zeldes (2012: 87), among others, therefore prefer the variable-corpus approach, which consists of comparing the $\mathscr{P}$–values and plot $\mathscr{P}$ and V in the P* plane for the largest sample size that they have in common. Given the wide variety of intensifiers in our data set and the relatively large proportion of hapax intensifiers overall, the token frequencies for the individual intensifiers are very small. For the top 15 intensifiers the largest shared sample size is only 12 tokens, and it would not be meaningful to compare the intensifiers at this token frequency. Still, to demonstrate how this approach could lead to different insights, we have recalculated the measures for the top 5 intensifiers at the largest common sample size of 60 (by taking random samples of 60 tokens for each intensifier) in Table 4.11.

Table 4.11. Frequency-based productivity measures for the top 5 intensifiers at N=60 in SoNaR-NL

	N	V	n1	$\mathscr{P}$	Hapax/type
rot	60	12	5	0.08	0.42
suf	60	37	29	0.48	0.78
dood	60	10	4	0.07	0.40
kapot	60	12	5	0.08	0.42
te pletter	60	22	18	0.30	0.82

We also replotted the global productivity graph to allow for a quick comparison between Figure 4.12 and Figure 4.13.

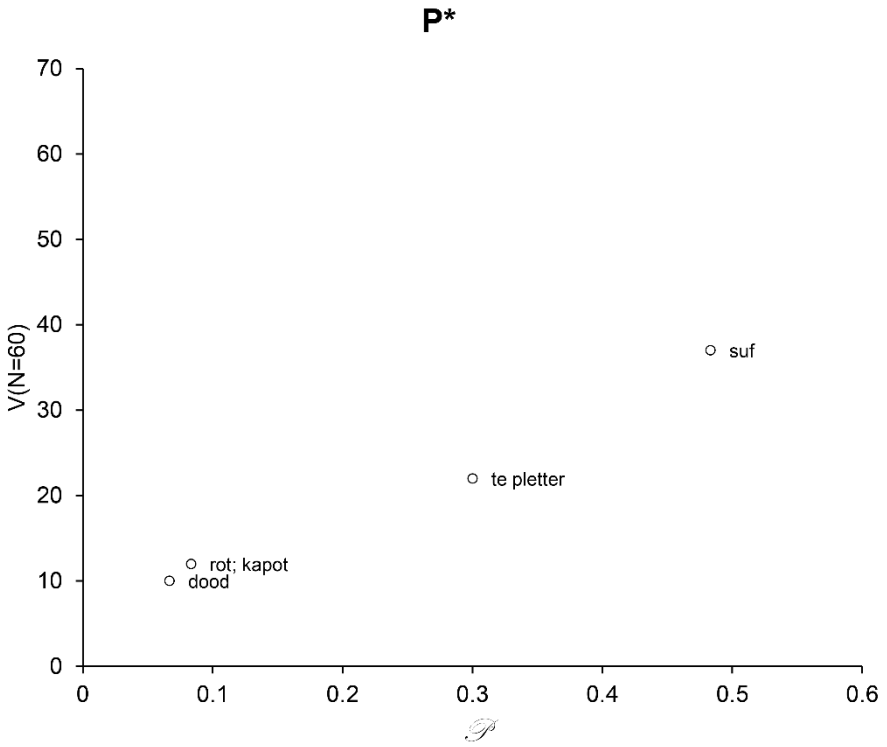


Figure 4.13. Global productivity (P*) for the top 5 intensifiers at N=60 in SoNaR-NL

As we can see, *suf* 'drowsy' is situated more to the bottom right in the smaller corpus size, compared to its position in Figure 4.12. The higher P^*-score in the smaller sample indicates that it is still more likely to occur with previously unattested types than in the larger corpus (Zeldes 2012: 88). Indeed, as the sample size increases, more types will be added to its collocational range; as a result, the collocational range becomes closer to being saturated and the potential productivity decreases. The same can be said for *dood* 'dead', although the position difference with Figure 4.12 is rather small: apparently *dood* is already much closer to having exhausted its productive potential at 60 tokens than *suf* 'drowsy'. Note that for *dood* 'dead', we have a larger hapax count in the smaller corpus. It does make sense that the chance of finding another attestation of a former hapax type, thus invalidating its hapax status, increases as the sample size increases. If we look at the actual hapax verbs in both sets, we see that *eten* 'to eat', *zich generen* 'to be embarrassed' and *vechten* 'to fight' have indeed lost their hapax status when the token frequency was increased from 60 to 116. At the same time, the chance of finding previously unattested hapaxes should also increase as more data are added to the sample. Indeed, the hapax verb *bidden* 'to pray' was added to the fold. The main point to take away from this is that we should ideally look at the actual hapax items in the data set to arrive at a nuanced picture of the changes in the collocational range. However, as it would take us too far to do this for every verb and intensifier in the set, the hapax *count* offers a more workable

alternative. *Kapot* 'broken' and *rot* 'rotten', although relatively close together in Figure 4.12 as well, now almost coincide into the same point. In the full data set, *rot* has slipped more to the left than *kapot*. That is, *rot* 'rotten' has gained proportionally fewer types (3 types in 94 tokens) than *kapot* 'broken' (4 types in 44 tokens), which may indicate that the $\mathscr{P}$-score for *rot* has decreased at a faster rate or, in other words, that it became increasingly difficult for *rot* to be extended to new types (cf. Zeldes 2012: 87-88, who observes the same development when comparing German word-formation patterns with *-lich* and *-bar*).

(b) Verbs

Although it is common practice in constructionist research to focus on the extensibility of the verb slot of a construction (cf. §2.1.2), the specific nature of the intensifying fake reflexive resultative construction and the attested covariation between the verb slot and the intensifier slot prompt us to also consider the productivity of the intensifier slot in subschemas with a lexically specified verb. Table 4.12 gives an overview of the frequency-based measures for the top 15 most frequently used verbs.

Table 4.12. Frequency-based productivity measures for the top 15 verbs in SoNaR-NL

	N	V	n1	$\mathscr{P}$	Hapax/type	H_{REL}
schrikken	223	19	8	0.04	0.42	0.514
zich ergeren	133	14	6	0.05	0.43	0.469
werken	112	22	5	0.04	0.23	0.638
zich schamen	80	4	0	0.00	0.00	0.296
lachen	70	17	7	0.10	0.41	0.544
zich vervelen	58	9	2	0.03	0.22	0.408
lopen	36	10	5	0.14	0.50	0.416
piekeren	22	3	2	0.09	0.67	0.087
betalen	18	4	1	0.06	0.25	0.232
zuipen	14	6	1	0.07	0.17	0.407
zoeken	14	8	6	0.43	0.75	0.421
eten	12	7	5	0.42	0.71	0.404
drinken	11	9	8	0.73	0.89	0.499
trainen	11	6	5	0.45	0.83	0.338
rennen	9	3	1	0.11	0.33	0.202

A quick glance at the table shows that the $\mathscr{P}$-values overall are rather low, which could be interpreted as verbs being generally less likely to be extended to new intensifiers than vice versa. However, if we take into account that the range of available intensifiers in general is smaller than the range of verbs, the chances of a verb appearing with a hapax intensifier are reasonably smaller than the chances of an intensifier appearing with a hapax verb. We could try to derive the potential productivity by turning to the hapax-

type ratio instead. This tells us that, for example, the verbs *schrikken* 'to be startled', *zich ergeren* 'to be annoyed' or *lachen* 'to laugh' have a lower degree of realised productivity than *werken* 'to work', but the former are more likely to expand their collocational range. It also shows that some of the verbs that are less frequent overall, like *drinken* 'to drink', *trainen* 'to train' or *zoeken* 'to search', allow for a greater deal of variation in their intensifier slot than the verbs most frequently attested in the construction. Interestingly, one of the top five of most frequent verbs appears to be completely unproductive, i.e. *zich schamen* 'to be embarrassed': it enters into conventional collocations with the 3 frequent intensifiers *rot* 'rotten', *dood* 'dead' and *kapot* 'broken' and it also enters into the exclusive collocation *zich de ogen uit het hoofd schamen* 'to embarrass oneself the eyes out of the head', but there are no creative combinations with this verb. Of course, the hapax-type ratio is easily inflated in case of overall low type frequency, much like $\mathscr{P}$ tends to be overestimated for low token frequent items (cf. supra). Consider the verb *piekeren* 'to worry': 2 out of 3 attested types are hapax legomena, which gives a high hapax-type ratio of 0.67, but this also means that the single remaining type accounts for 20 of the 22 tokens. This extreme skewness is evident in the low relative entropy score, but we already pointed out in the previous paragraph that the interpretation of relative entropy with respect to productivity is not always straightforward and that H_{rel} does not really add a lot of information that cannot be gleaned from other measures in the productivity complex. Again, it appears that the verbs that have the highest degree of productivity, based on the other frequency-based measures, also have relatively high entropy scores, e.g. *werken* 'to work', *schrikken* 'to be startled', *lachen* 'to laugh', compared to verbs like *piekeren* 'to worry' or *betalen* 'to pay', the distributions of which are dominated by one intensifier, viz. *suf* 'drowsy' and *blauw* 'blue', respectively. Overall, it appears that the distribution is less informative for the verb perspective than for the intensifiers. With the exception of some specific verbs, the verbs generally do not have such pronounced preferences and show a more balanced distribution across the intensifiers they co-occur with. This was already obvious from the covarying collexeme analysis as well, see Table 4.6 above. The ΔP-values for the strong collocations were usually higher for the intensifier-to-verb attraction, indicating that the asymmetric association is primarily dependent on the combinatorial behaviour of the intensifier. The realised productivity and the potential productivity of the verbs are visually represented in the global productivity graph in Figure 4.14.

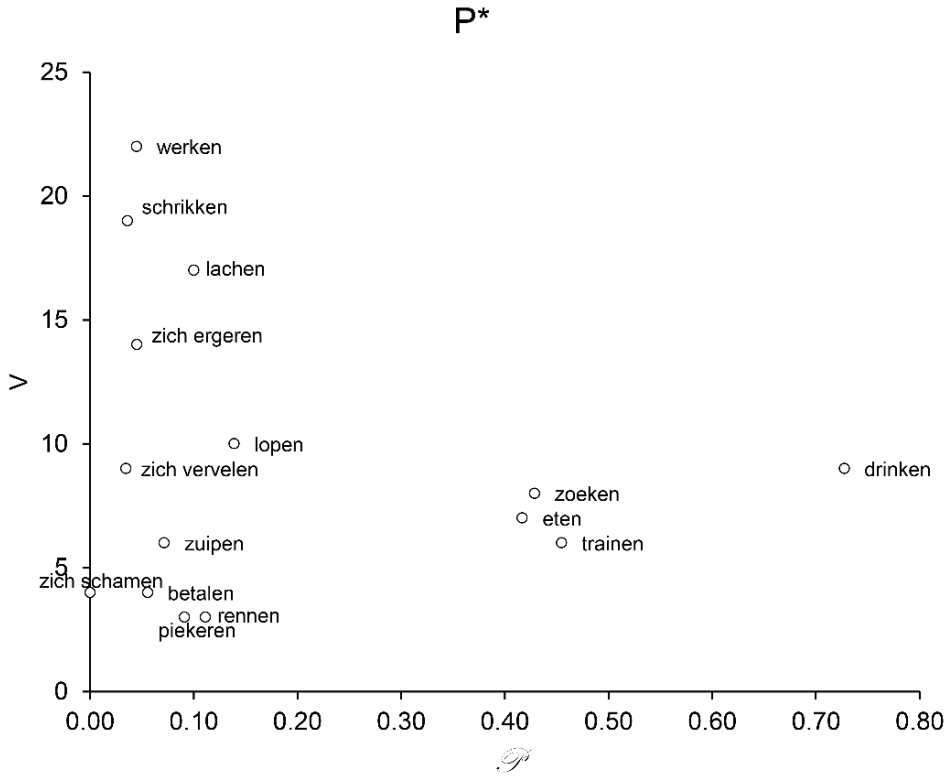


Figure 4.14. Global productivity (P*) of the top 15 verbs in SoNaR-NL

While the intensifiers were mostly situated in the lower part of the graph, spread out vertically along the X-axis, the bulk of the verbs are spread out along the Y-axis in the left hand part of the graph. We can group the verbs together in global productivity clusters in much the same way as we did for the intensifiers. Starting at the bottom left hand side of the graph, we see the verbs *zich schamen*, *betalen*, *rennen*, *piekeren* and *zuipen*, which were found to co-occur with a small set of types (i.e. low degree of realised productivity) and are not likely to be extended to many more types, based on their low proportion of hapaxes. This low probability of extensibility also applies to the verbs *zich vervelen* and *lopen*, but unlike the previous group, these verbs must have been able to attract new intensifiers in the past, as they have accumulated a number of intensifiers in their collocational range. Moving further along the Y-axis, we could then group together the verbs *werken*, *schrikken*, *lachen* and *zich ergeren*, which have the highest degree of realised productivity. In the middle of the plane, we find a small group of *zoeken*, *eten* and *trainen*, which appear to have a higher degree of potential productivity, while also displaying a certain type frequency. A real outlier here is *drinken*, in that it appears to have a much higher probability of occurring with a previously unattested intensifier than all other verbs. However, given its low token frequency overall, it is hard to really substantiate this claim. Once more, it would be more viable to compare the productivity of the verbs at their largest shared sample size but for the top 15 verbs, this common

token size is only at 9 tokens. Table 4.13 shows the recalculations for just the top 5 verbs at the shared sample size of 70 tokens.

Table 4.13. Frequency-based productivity measures for the top 5 verbs at N=70 in SoNaR-NL

	N	V	N1	$\mathscr{P}$	Hapax/type
schrikken	70	13	6	0.09	0.46
zich ergeren	70	9	3	0.04	0.33
werken	70	17	6	0.09	0.35
zich schamen	70	4	0	0.00	0.00
lachen	70	17	7	0.10	0.41

A comparison of the numbers in Table 4.11 and Table 4.13 suggests that the interpretation of the $\mathscr{P}$-score should be taken with a pinch of salt. *Zich ergeren* 'to be annoyed' has a lower $\mathscr{P}$-value than *werken* 'to work' and *schrikken* 'to be startled' at 70 tokens, which would suggest that *zich ergeren* is less likely to be extended to new types. However, when *zich ergeren* doubles in token frequency, its type frequency is increased by 5 types (55% of the original type frequency), whereas *schrikken,* which triples in token frequency, only gains 6 types (46% of the original type frequency). It therefore seems that *schrikken* 'to be startled' was already closer to saturating its productive potential at 70 tokens than *zich ergeren* 'to be annoyed' was, in contrast to what the $\mathscr{P}$-value would suggest.

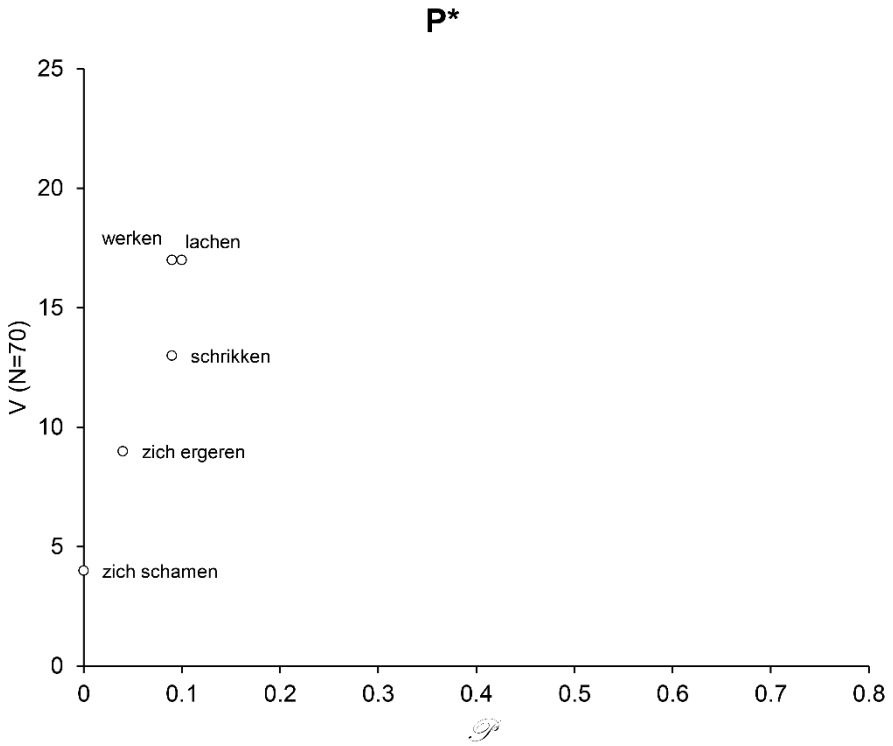


Figure 4.15. Global productivity (P*) of the top 5 verbs at N=70 in SoNaR-NL

To summarise, if we take together all the indivual intensifiers and verbs, we could, broadly speaking, distinguish four large "productivity classes". First, we have a number

of items with a relatively high overall degree of productivity, in that they have moderate to relatively high scores for both realised and potential productivity (high V, high $\mathscr{P}$), e.g. *een slag in de rondte* 'a punch around', *een ongeluk* 'an accident', *zoeken* 'to search', *drinken* 'to drink'. A second group consists of those items with a degree of realised productivity containing traces of past productivity, but the productive potential of which is nearing exhaustion (high V, low $\mathscr{P}$), e.g. *kapot* 'broken', *werken* 'to work', *schrikken* 'to be startled'. In the third group we find the verbs and intensifiers that have a limited collocational range at present, but which do show some potential productivity (low V, high $\mathscr{P}$), e.g. *wezenloos* 'blank/vacant' (no verb examples). Finally, there is a group of "unproductive" items (low V, low $\mathscr{P}$), e.g. *een hoedje* 'a little hat', *wild* 'wild', *zich schamen* 'to be embarrassed', *piekeren* 'to worry'. We will expand on the network implications of these productivity classes by showing how subschemas that display varying degrees of productivity, should be represented in the overall structure of the network in section §4.4.1 below.

One aspect that was not taken into account thus far concerns the semantics of the items that fill the empty slots in the subschemas. That is, while type frequency can give us an idea of the lexical diversity of a construction by counting how *many* different types occur in the construction, it does not tell us what these items are and how they are semantically related. In addition, the hapax counts only tell us how many one-offs occur in the construction, but not how truly novel these really are (cf. supra). Earlier in this chapter, the hierarchical cluster analyses revealed that the collocational behaviour or collocational patterns in which the verbs and intensifiers interact often do not appear to be motivated by the original lexical semantics in as far as semantically related verbs/intensifiers do not necessarily show the same collocational preferences and, conversely, semantically unrelated verbs/intensifiers can enter into similar collocational patterns. That is not to say, of course, that semantics should be disregarded completely, quite the contrary. In Chapter 2, §2.1.2, we mentioned several studies that have suggested that the productivity of constructions can be limited if the construction is subject to semantic constraints. Indeed, it feels rather counterintuitive to say that *dood* 'dead' or *rot* 'rotten' have a lower degree of productivity compared to a lexically-specific intensifier like *de longen uit het lijf* 'the lungs out of the body' – although this is what the frequency-based measures seem to suggest. In the next paragraph we will add further nuance to the productivity scores by also taking into account the semantics of the verbs that co-occur with some of these intensifiers. The next paragraph will complement the frequency-based productivity analysis by adding semantic coherence as one of the dimensions of productivity in our model.

4.3.1.2 A constructional model of productivity

The first theoretical model of productivity that attributes a central place to the semantic aspect of productivity is the constructional model by Barðdal (2008), which assumes that the productivity of a (sub)schema can be predicted by the inverse correlation between type frequency and semantic coherence, the implications of which were represented by the cline in Figure 4.9 above. Concretely, a subschema which can be instantiated by a high number of types does not require its types to display a great degree of semantic coherence in order to be productive. The inverse correlation implies that subschemas that are not very type frequent can also be productive if their types belong to a coherent semantic domain. In that case, the subschema is said to be productive within that delimited semantic domain. At the most extreme point of this cline, a subschema is instantiated by just one type and is therefore maximally coherent. This brings us back to the issue of high token frequency items. In the traditional literature on morphological productivity, it was assumed that the existence of types that were extremely token frequent detracts from the productivity of a pattern. The reasoning behind this assumption is that these highly token frequent types are stored as prefabricated chunks rather than as on-the-fly instantiations of a productive pattern (cf. Ch2, §2.1.2). If we look at the data presented in the previous section, we find several instances for which this prediction appears to be borne out. There are several intensifiers and verbs that enter into (near-)exclusive combinations or collocations and can therefore be said to be "unproductive", e.g. *zich* ***groen en geel*** *ergeren, zich suf* ***piekeren****, zich* ***een hoedje*** *schrikken, zich* ***het vuur uit de sloffen*** *lopen, zich* ***wild*** *ergeren/schrikken*. It would seem that the extension of these particular intensifiers or verbs to new types is blocked, or at least highly constrained, by the existence of such strong collocations. Although Barðdal does not deny that this is a possible effect of high token frequency, she suggests that it is not the only possible scenario: alternatively, highly frequent types may also *encourage* the extension to new types. In this scenario, the single token frequent type comes to serve as a model that can trigger new extensions via analogy. What is crucial in her proposal is that these new types are necessarily semantically related to the high-frequency type. The importance of semantics has been demonstrated in studies focusing on variation in idioms as well (Gibbs et al. 1989, Erman & Warren 2000). In a more recent monograph, Zeschel (2012) further substantiates this scenario by demonstrating how low-scope productive schemas may emerge out of conventionalised, fixed expressions, and eventually may even give rise to creative extensions beyond the specific semantic scope. This interaction between frequency and semantics in determining (syntactic) productivity is also confirmed by an experiment by Suttle & Goldberg (2011), who investigated the contribution of three factors, viz. type frequency, variability and similarity (see Ch2, §2.1.2). They found that both type frequency and variability had strong independent effects, but they were also part of an

interaction in that the effect of type frequency was found to be stronger in cases of high variability (i.e. low coherence).

(a) Semantic coherence in the intensifying fake reflexive resultative construction

If we start out by taking a look at the macro-level construction, i.e. the schematic pattern [SUBJ V REFL INT], we see that both the verb slot and the intensifier slot are highly type frequent and that the types can be recruited from a number of unrelated semantic domains (cf. §4.1.1 supra). Once more we obtain very different results if we drop down to the subschema level, especially if we look at the combinatorics of some specific intensifiers. It appears that the role of semantic coherence is less evident in the collocational behaviour of the verbs. Most of the frequently used verbs can occur with a wide variety of intensifiers from different syntactic and semantic categories. Still, we do find some verbs for which several intensifiers occurring in the intensifier slot form a semantically coherent group. Some of the intensifiers occurring with the verbs *lopen* 'to run' and its synonyms *rennen* 'to run' and *hollen* 'to run', for example, are NP+PP intensifiers involving legs or footwear, e.g. *de benen uit het lijf* 'the legs out of the body', *de benen uit het gat* 'the legs out of the butt', *de blaren op de voeten* 'the blisters on the feet', *het vuur uit de sloffen* 'the fire out of the slippers', see (166) and (167) infra. In such cases, there appears to be some kind of mutual attraction between the semantics of the verb and the semantics of the intensifier, but it will become clear that this coherence is mainly explained by restrictions imposed by the intensifier, rather than the verb. In fact, *lopen* 'to run' readily pairs up with lexically unspecified intensifiers like *rot* 'rotten', *te pletter* 'to smithereens' or *suf* 'drowsy' as well.

(151) Buurvrouw A. Pijnaker van Allicht Cronjé **liep zich** eind 2001 "**de benen uit mijn gat**". (SoNaR-NL)
[...] ran herself late 2001 the legs out of my butt
'Neighbour A. Pijnaker van Allicht Cronjé ran her socks off at the end of the year 2001.'

(152) Joop komt één uur per dag uit bed. Hij ziet dagelijks hoe de meisjes **zich rot lopen**. (SoNaR-NL)
[...] the girls themselves rotten run
'Joop gets out of bed one hour a day. He sees the girls running their socks off every day.'

All in all, verbs generally do not seem to impose important semantic constraints on the open intensifier slot - or, in other words, the intensifier slot is not very sensitive to the principle of semantic coherence – in present-day Netherlandic Dutch.

Of course, there are also several intensifiers that can be used with a wide range of verbs from several of the semantic classes that were established earlier. Among the top ten of the most frequent intensifiers, we see such "all-round" behaviour for *rot* 'rotten', *dood* 'dead', *suf* 'drowsy', *kapot* 'broken', *te pletter* 'to smithereens' and *een ongeluk* 'an accident'. That is not to say that these intensifiers do not show collocational *preferences*, just that

they are not really constrained to any specific semantic class. *Een slag in de rondte* 'a punch around', and by extension *n° slagen in de rondte* 'n° punches around' are also semantically very flexible intensifiers in that they can occur with a wide range of verbs, viz. physical activity verbs, communication verbs, consumption verbs or other general activities, see (153) to (157).

(153) Hij **tekende, schilderde en boetseerde zich een slag in de rondte**, maar kwam er niet uit. (SoNaR-NL)
he drew, painted and moulded himself a punch around [...]
'He drew, painted and moulded piece after piece, but he couldn't figure it out.'

(154) Ackroyd heeft **zich een slag in de rondte gezocht** naar bronnen. (SoNaR-NL)
ackroyd has himself a punch around searched for sources
'Ackroyd has searched everywhere to find sources.'

(155) Cultuur, overschatten we de betekenis niet? We **discussiëren ons** hier **een slag in de rondte**. (SoNaR-NL)
[...] we discuss ourselves here a punch around
'Culture, are we not overestimating the meaning of it? We are having endless discussions about it.'

(156) Man, ik **werk me** elk weekend **een slag in de rondte** voor mijn programma's Nachtsuite en Lunchroom (SoNaR-NL)
man, I work myself every weekend a punch around [...]
'Man, I'm working my butt off every weekend for my shows Nachtsuite and Lunchroom.'

(157) Binnen de kortste keren begint hij kamerbreed al zijn afspraken te schenden - hij **zuipt en rookt zich een slag in de rondte**. (SoNaR-NL)
[...] boozes and smokes himself a punch around
'In no time he starts violating all of his arrangements – he drinks and smokes profusely.'

It appears that all verbal collocates of *een slag/slagen in de rondte* 'a punch/punches around' must denote some kind of *activity*: the experience verbs like *lachen* 'to laugh', *schrikken* 'to be startled', *zich ergeren* 'to be annoyed', etc. are remarkably absent from the collocational range.[46] There is no easy explanation as to why *een slag in de rondte* 'a punch around' has such a strong dispreference for experience verbs, although it may have something to do with the element of *slag* 'punch' in the intensifier, which could invoke some kind of activity frame.

There are several examples in which the verb and intensifier enter into a semantic relation, in the sense that the semantic constraints appear to be motivated by the original lexical semantics of the intensifier. Take, for example, *de longen uit het lijf* 'the lungs out of

[46] Full list of verbs with *een slag/slagen in de rondte* 'a punch/punches around': *associëren, boetseren, borduren, dirigeren, discussiëren, drinken, duiken, eten, kletsen, knokken, neuken, organiseren, overleggen, rijden, roken, trainen, transformeren, vergaderen, vrijen, werken, werven, zoeken, zuipen* (for English translations, see the translation list at the beginning of the thesis)

the body', which naturally occurs with verbs that denote an activity for which one needs to rely (in some way or another) on one's lung capacity. Looking at the verbal collocates, we see that this semantic restriction does allow for some freedom of interpretation. First of all, there is a more or less coherent group of verbs that involve sound emission or air expulsion, viz. *blazen* 'to blow (an instrument)', *gillen* 'to screech', *schreeuwen* 'to scream' and *zingen* 'to sing', see (158) and (159).

(158) Aan de rand van het centrum **zingen en blazen** musici als Angie Stone en Wynton Marsalis **zich** twee weekenden achter elkaar **de longen uit het lijf**. (SoNaR-NL)
[...] sing and blow musicians [...] themselves [...] the lungs out of the body
'At the edge of the city centre, musicians like Angie Stone and Wynton Marsalis are singing and playing their hearts out two weekends in a row.'

(159) Jonathan Davis **gilde zich de longen uit het tengere lijf**. (SoNaR-NL)
jonathan davis screeched himself the lungs out of the frail body
'Jonathan Davis screeched his head off.'

Another group of verbs are the verbs denoting a heavy physical exercise, viz. *fietsen* 'to cycle', *knokken* 'to fight', *lopen* 'to run' and *zwemmen* 'to swim', see (160).

(160) Frank Heldoorn (34) **zwemt, fietst en loopt zich** al jaren **de longen uit het lijf**, maar mist de gunst van het grote publiek. (SoNaR-NL)
frank heldoorn swims, cycles and runs himself already years the lungs out of the body [...]
'Frank Heldoorn (34) has been swimming, cycling and running his butt off for years, but he does not find favour with the general public.'

These verbs are intuitively quite different from the previous group, but both verb classes do share the general idea that a large lung capacity is important in performing these activities. In the next section, we will see how these restrictions on the verb slot can be dealt with in the structure of the constructional network.

A very similar semantic constraint is imposed by the intensifier *uit de naad* 'out of the seam', which actually shares some specific collocates with *de longen uit het lijf* 'the lungs out of the body'. *Uit de naad* prefers verbs that express some kind of physical effort, viz. *lopen* 'to run', *rijden* 'to ride', *swingen* 'to swing', *trainen* 'to train' and *werken* 'to work'. By far the most frequent collocate is *werken* 'to work', which was absent in the collocational range of *de longen uit het lijf* 'the lungs out of the body'. We also find one occurrence with the noise emission verb *zingen* 'to sing', but it is coordinated with the physical effort verb *swingen* 'to swing' in (163). In the case of *uit de naad* 'out of the seam', the semantic constraint is not immediately explained by a specific lexical element of the intensifier but it may have something to do with the image of someone bursting out of the seams of their clothes as a result of performing an activity with a certain intensity, see (161) to (163).

(161) Het is toch stom dat je **je** eerst **uit de naad rijdt** en vervolgens de trui laat schieten. (SoNaR-NL)

[...] that you yourself first out of the seam ride [...]
'It is silly to first ride your butt off in each race only to pass over the opportunity of actually bringing home the jersey.'

(162) Ik **werkte me uit de naad**, op zeer onregelmatige tijdstippen. (SoNaR-NL)
I worked myself out of the seam [...]
'I worked very hard at very irregular hours.'

(163) Geboren show-mens Simone Kleinsma (45) **zingt en swingt zich uit de naad** in de musical Mamma Mia! (SoNaR-NL)
[...] sings and swings herself out of the seam [...]
'Born showstopper Simone Kleinsma (45) sings and swings her butt off in the musical Mamma Mia!'

Another example are the less frequent intensifiers *de blaren op de tong* 'the blisters on the tongue' or *de blaren op het verhemelte* 'the blisters on the palate', which co-occur with activities that involve the use of one's mouth. This constraint is less freely interpretable than the one for *de longen uit het lijf* 'the lungs out of the body', but we do see *some* variety in the verb types, viz. *praten* 'to talk', *kletsen* 'to chatter' and *onderhandelen* 'to negotiate'.

(164) Generaties theologen hebben **zich de blaren op het verhemelte gekletst** om hun god vrij te pleiten. (SoNaR-NL)
generations theologians have themselves the blisters on the palate chattered [...]
'Generations of theologians have talked their heads off to exonerate their god.'

(165) En dat terwijl wij **ons de blaren op de tong** hebben **onderhandeld** om voor volgend jaar op de nullijn te blijven. (SoNaR-NL)
and that while we ourselves the blisters on the tongue have negotiated [...]
'Even though we have negotiated like crazy to stay at the neutral line next year.'

Even more strict are the constraints imposed by the intensifiers *het vuur uit de sloffen* 'the fire out of the slippers' and *de benen PREP het lijf* 'the legs out of the body'. In the SoNaR-NL data set, *het vuur uit de sloffen* is exclusively found with the verb *lopen* 'to run', see (166), and *de benen uit het lijf* occurs with *lopen* and its synonym *rennen* 'to run'. It would appear that the verb slot is lexically constrained, meaning that it only allows for one specific verb (and its nearly perfect synonyms) (cf. Zeschel 2012: 7).

(166) Dean **loopt zich het vuur uit de sloffen** om stemmen te werven. (SoNaR-NL)
dean runs himself the fire out of the slippers [...]
'Dean runs his socks off to canvas for more votes.'

(167) De Serviër **loopt zich de benen onder het lijf vandaan** , biedt zichzelf in alle gretigheid zelfs op het middenveld aan. (SoNaR-NL)
the serbian runs himself the legs under the body off [...]
'The Serbian runs his socks off, even presenting himself eagerly in the centrefield.'

(168) Een bemanning die de schipper en de officieren enerzijds van alles wat misging de schuld gaf, maar anderzijds voor een oorlam **zich** ook weer **de benen uit het lijf rende**. (SoNaR-NL)

[...] itself also again the legs out of the body ran
'A crew that, on the one hand, blamed the shipmaster and his officers for everything that went wrong but, on the other hand, they would scurry around the ship for a drink.'

The collocational constraints look to be especially prominent for the NP+PP intensifiers, which are lexically much more specific and "contentful" than any of the other categories. This may also explain why the NP+PP intensifiers are generally not found to co-occur with the inherently reflexive verbs: these verbs express some kind of emotional or cognitive experience that is hard to reconcile with the original meaning of the intensifiers in question. In as far as the original lexical semantics of these intensifiers impose a restriction on the verbal range, the semantic constraints can be very specific and may even give the impression of being rather "ad hoc". That is, the lexical elements in *het vuur uit de sloffen* 'the fire out of the slippers' and *de benen uit het lijf* 'the legs out of the body' come together in a very specific lexical constraint, viz. *lopen* (or synonyms) 'to run'. The same could be said for the more global constraint of "lung capacity" in case of *de longen uit het lijf* 'the lungs out of the body', subsuming the semantic classes of physical exercise verbs and noise emission verbs, two classes which may not be thought of as semantically related outside of this particular context. Similarly, in case of the intensifiers *de blaren op de tong* or *de blaren op het verhemelte* 'the blisters on the tongue/palate', the semantic constraint of [+mouth] groups together a very specific set of verbs. The adjectival and some prepositional and nominal intensifiers, on the other hand, are usually more "all-round" intensifiers, in that they do not seem to impose any obvious semantic constraints on the items they co-occur with. Many intensifiers have shed their original lexical semantics and have come to function as pure, unconstrained intensifiers in this construction.

The historical development of these intensifiers will be investigated in Chapter 5, but the synchronic data already contain some indications that at least one of the all-round intensifiers used to be considerably more limited in its use. *Suf* 'drowsy' is currently the second most frequent intensifier and outperforms all other intensifiers in terms of collocational range. Nonetheless, among the 61 verb types, there is a remarkably coherent group of verbs that express some kind of mental activity.

(169) Advocaat **prakkiseert zich suf** over een nieuwe rol voor Frank de Boer. (SoNaR-NL)
advocaat broods himself drowsy [...]
'Advocaat is brooding hard over a new role for Frank de Boer.'

(170) Twaalf medewerkers en redacteuren van Cicero **piekerden zich suf** om drie titels te selecteren. (SoNaR-NL)
twelve employees and editors of cicero worried themselves drowsy [...]
'Twelve employees and editors of Cicero worried a lot about picking three titles.'

As this category appears to be related to the intensifier's original lexical meaning, in much the same way as was observed for the NP+PP intensifiers, it is plausible that *suf*

'drowsy' started out as an intensifier that imposed the specific semantic constraint of [+mental activity] on its verbal collocates. Over time, *suf* appears to have developed into a semantically bleached, all-round intensifier that is able to co-occur with a large variety of different verbs.

(171) Op mijn hotelkamer **zapte ik me** woensdagavond helemaal **suf** om de wedstrijd live in beeld te krijgen. (SoNaR-NL)
in my hotel room zapped I myself wednesday evening totally drowsy [...]
'I flipped through the channels like crazy in my hotel room on Wednesday evening so I could watch the game live.'

(172) Het zijn mensen die **zich suf ergeren** aan modieuze pleidooien voor orde, gezag en een goed pak slaag. (SoNaR-NL)
they are people who themselves drowsy worry [...]
'They are people that worry a lot about fancy pleas for order, authority and a good thrashing.'

Of course, we need diachronic data to confirm whether this is actually what happened, which is why Chapter 5 will pay special attention to the changes within the collocational patterns and the possibility of relaxing collocational constraints in recent history. In the same vein, some of the semantically restricted intensifiers that were discussed earlier may also gradually be developing a more all-round intensifier function over time. Judging by example (173) from the Internet, it appears that – at least for some speakers – the semantic constraints imposed by *de longen uit het lijf* 'the lungs out of the body' have somewhat relaxed. At the same time, such an example could be interpreted as a deliberate override of the semantic restrictions in order to create a special rhetorical effect or to draw the attention of the hearer/reader (see also §4.2.1 supra).

(173) Ik **schrok me de longen uit het lijf** pff. (www.dumpert.nl 2013)
I startle myself the lungs out of the body pff
'I was very started, pfff.'

(b) A first step towards refining and operationalising the constructional model

Based on the data we have presented thus far, both type frequency and semantic coherence are important factors when it comes to determining the degree of productivity of the subschemas. To better visualise the different degrees of productivity, we could try putting some of the intensifiers on the productivity cline suggested by Barðdal (2008). At first blush, there are a number of intensifiers that could be put on the cline quite straightforwardly. Given the extremely high type frequency and semantic diversity of the verbs occurring with *suf* 'drowsy', we could put that intensifier all the way at the top of the cline. The intensifier *de longen uit het lijf* 'the lungs out of the body' could be positioned somewhere in the lower half of the cline: it has a much lower type frequency than *suf* 'drowsy' and it clearly imposes semantic constraints on the verbs it occurs with. The

verbal collocates of *de longen uit het lijf* 'the lungs out of the body' belong to two distinct (but apparently related) semantic classes, viz. the verbs denoting a heavy physical activity and the verbs expressing some kind of air expulsion or noise emission. Slightly more to the bottom right, we can put *uit de naad* 'out of the seam'. It has a lower type frequency than *de longen uit het lijf* 'the lungs out of the body' and appears to be stricter with respect to the semantics of its collocates. Although the occasional noise emission verb is perhaps not excluded (see example (163) with *zingen* 'to sing'), it has an outspoken preference for verbs denoting some kind of physical effort. Even more to the bottom right, we could posit the intensifier *de blaren op de tong* 'the blisters on the tongue', with 3 verb types involving some kind of communicative activity. Finally, given that *het vuur uit de sloffen* 'the fire out of the slippers' is exclusively combined with the verb *lopen* 'to run' in the SoNaR-NL data set, we would have to put it at the extreme bottom end of the cline. Without claiming that these are their exact positions on the cline, we propose a possible representation of the different degrees of productivity in Figure 4.16.

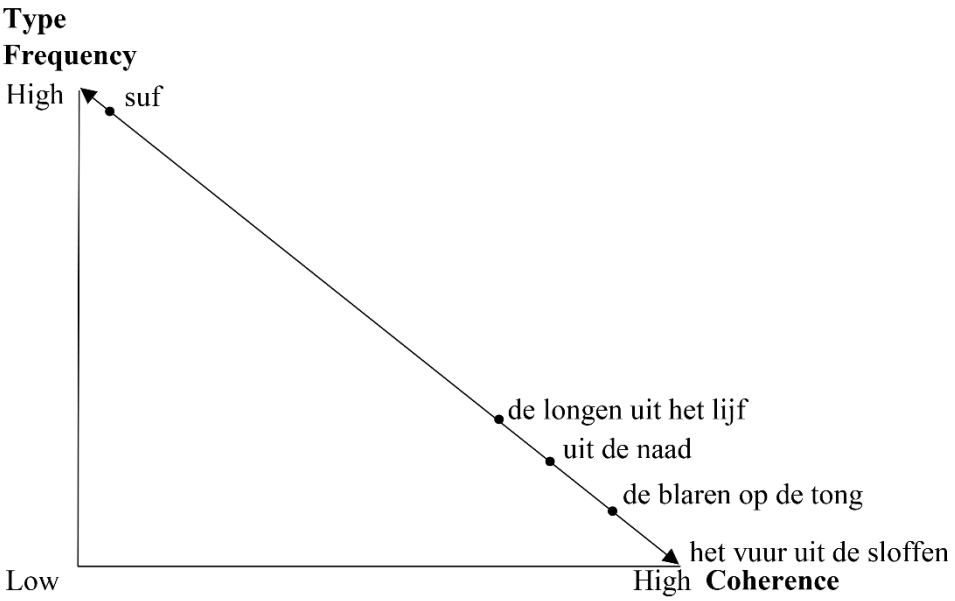


Figure 4.16. Position of some intensifiers on the productivity cline, adapted from Barðdal (2008: 35)

Nevertheless, we may wonder what to do with some of the other all-round intensifiers. In terms of type frequency, the vertical position of the intensifiers *dood* 'dead' (13 types) and *rot* 'rotten' (14 types) should be much closer to *de longen uit het lijf* 'the lungs out of the body' (8 types) than to *suf* 'drowsy' (61 types). Given the absence of obvious semantic coherence among their verb types, however, they should be more or less at the same position on the X-axis as *suf* 'drowsy'. As a result, it appears that they can no longer be fit straightforwardly onto the cline. It gets even more complicated when we want to add the intensifier *een slag in de rondte* 'a punch around'. With 22 types it has a higher type frequency than some of the all-round intensifiers, but it is semantically more constrained because it appears to have a dispreference for experience verbs. This is clearly not what

we would expect given the inverse correlation that is at the basis of the productivity model. Another curious example is found in *wezenloos* 'blank/vacant'. We find that *wezenloos* only has three more types than *uit de naad* 'out of the seam', but *uit de naad* is semantically constrained (cf. supra), whereas *wezenloos* belongs to the all-round intensifiers, as illustrated by the examples below.[47]

(174) De Fransen **schrokken zich wezenloos** toen Moskou verlaten bleek. (SoNaR-NL)
the french startled themselves vacant [...]
'The French were very startled when Moscow turned out to be deserted.'

(175) De NT'er **communiceert zich wezenloos**, maar wat presteert hij eigenlijk? (SoNaR-NL)
the nt'er communicates himself vacant [...]
'The NT'er is communicating all over the place, but what does he actually accomplish?'

(176) We hebben **ons wezenloos gezocht** waar het vuur zit, maar we kunnen het niet vinden. (SoNaR-NL)
we have ourselves vacant searched [...]
'We have looked everywhere for the origin of the fire, but we cannot find it.'

It was pointed out that subschemas with a "sufficient type frequency" do not need any semantic coherence to be productive, but this raises the question of what counts as "sufficient"; that is, why is 6 types not enough for *uit de naad* 'out of the seam' to be extended beyond its specific semantic domain, whereas 9 types appears to be enough to warrant the extensibility of *wezenloos* 'blank/vacant'? To be sure, there is no universal type frequency threshold that applies to all intensifiers and it may be necessary to approach the productivity of some subschemas on an item-by-item basis. For one, the contrast between *wezenloos* 'blank/vacant' and *uit de naad* 'out of the seam' may have something to do with the fact that one physical effort verb, i.e. *werken* 'to work', accounts for about 70% of all tokens of *uit de naad* 'out of the seam' (cf. supra). While *wezenloos* 'blank/vacant' also has a clear collocational preference, viz. for the verb *schrikken* 'to be startled', the distribution is slightly less skewed. That is, in addition to the mere type frequency, the type-token distribution may also play a role in the extensibility of the construction (cf. the discussion in §4.3.1.1).

Another peculiarity is the intensifier *blauw* 'blue', which has a rather limited type frequency of 7 types that barely show any semantic coherence at all. *Blauw* 'blue' is part of two strong collocations *zich blauw ergeren* 'to annoy oneself blue' (9 instances) and *zich blauw betalen* 'to pay oneself blue' (12 instances), which belong to very different semantic classes. Outside of these collocations, *blauw* 'blue' is found with 5 other infrequent verb types. If these other verb types were analogical extensions based on the strong model collocations, we would expect them to be semantically similar to the models. It appears

[47] Full list of verbs with *wezenloos* 'blank/vacant': *bellen, bezuinigen, communiceren, kopen, schrikken, sms'en, zich ergeren, zich verheugen, zoeken* (for English translations, see the translation list at the beginning of the thesis)

that most of the creative, one-off instances of *blauw* feature verbs that are not directly related to either of the token frequent verbs and are scattered across semantic space. We could argue that there is a certain logical connection between *zich ergeren* 'to be annoyed' and *klagen* 'to complain', but *bellen* 'to call', *lezen* 'to read', *piekeren* 'to worry' and *schrijven* 'to write' are harder to account for in the scenario of analogy. At any rate, based on its current distribution, we would expect *blauw* 'blue' to hardly show any signs of productivity at all.

(177) Van de begroting wordt 38 miljoen euro betaald door de financiële bedrijven die gecontroleerd worden. Zij **klagen zich blauw** over de kostenstijging. (SoNaR-NL)
[...] they complain themselves blue about the cost increase
'38 million euros of the budget is paid by financial companies that are under external supervision. They are complaining a lot about the cost increase.'

(178) We zitten **ons blauw** te **bellen** en we sturen bemiddelaars en gezanten omdat we denken dat het op die manier mogelijk is partijen bij elkaar te brengen. (SoNaR-NL)
we sit ourselves blue to call [...]
'We are making a lot of phone calls and sending out intermediaries and envoys because we think that this will allow us to bring the parties together.'

While the results so far confirm that there is undeniably some kind of interaction between type frequency and semantic coherence, their supposed *inverse correlation* makes certain predictions about the (non-)productivity of some intensifiers that do not always appear to be borne out when we look at the actual data. Granted, the specific expressive and intensifying nature of the construction under investigation may provide some explanation as to why the model does not provide a perfect fit for our data. It was already argued in Chapter 2, §2.3, that there are certain forces at work in the domain of intensification, such as the need for extravagance, that do not apply to more neutral descriptive constructions. The current version of the productivity model may have a much stronger explanatory force for the productivity of "regular" argument structure constructions. Barðdal (2008) herself discusses several case studies that illustrate how the model is to be applied to actual data and convincingly demonstrates that the predictions of the model are borne out for certain case and argument structure constructions in Icelandic. In the first case study, for example, she investigates the constructional preferences of a set of 107 verbs that were recently borrowed from English into Icelandic. The Nom-Acc construction has the highest type frequency and the highest semantic scope (i.e. lowest coherence) and should therefore be situated all the way at the top of the cline. The Nom-Dat construction has an intermediate frequency and a considerably lower semantic scope (i.e. higher coherence) than the Nom-Acc construction, which would correspond to a position halfway along the cline. The Nom-Gen construction is of relatively low type frequency, compared to the other subconstructions, and is made up of some lexically idiosyncratic phrases and some verb-specific islands that show very little coherence. With both low type frequency and low coherence, the Nom-Gen

construction is situated far away from the cline and should therefore display no signs (or only sporadic traces) of productivity. The productivity of the subconstructions of the Nominative subject construction appears to conform to the predictions made by the inverse correlation between type frequency and semantic coherence. Of the 107 new verbs, the majority select for the Nom-Acc pattern (the pattern highest on the cline), the Nom-Dat attracts some verbs belonging to the same semantic domains as the already attested verbs and none have selected the Nom-Gen pattern (the pattern that is not on the cline). In another case study, she compares the extensibility of the transfer construction (a Nom-Acc subpattern) and the caused-motion construction (the most productive Nom-Dat subconstruction) to new verbs of communication like *emeila* 'to email'. Concretely, she finds that the productivity of the transfer construction with these new verbs of communication is primarily based on analogy with one existing verb of sending, *senda* 'to send' (at the bottom extreme of the cline). The caused-motion construction was originally not used with verbs of sending at all, but it has been gaining some productivity recently because it has a higher type frequency than the transfer subconstruction and is less bound by semantic coherence. Even so, we would suggest to slightly relax the rigour of the inverse correlation by allowing for some more flexibility with respect to the position in relation to the cline. Our data have shown that some intensifiers can still display a certain degree of productivity even if they appear to be positioned quite far away from the cline (i.e. relatively low type frequency and coherence). This also brought us to the question of what counts as "high" and "low" type frequency and coherence. While most frequency-based studies are aware of the problematic nature of frequency thresholds, the question largely remains unresolved. A possible solution, as has been proposed by Baayen & Lieber (1991), among others, is to adopt a more relative interpretation. Rather than using the "exact" position of a category to predict its productivity (or non-productivity), we could compare the positions of multiple categories to one another in order to find out which has a higher degree of productivity, in much the same way as we did for the global productivity in the previous section.

Even if we were to slightly adjust or relax the implications of the productivity model, its main drawback is that, in its current state, it cannot really be operationalised in a systematic fashion. While we can easily add the type frequency values on the Y-axis, semantic coherence is a more difficult notion to operationalise because meaning is not directly observable, at least not in the same way as formal properties are. We would need to determine the criteria on the basis of which semantic coherence can be determined (i.e. how are verb classes defined) and find a way to quantify semantic coherence. Linguistic meaning is often incorporated in corpus studies through manual annotation of semantic features (abstractness, animacy...) or broadly defined semantic classes (as the ones established in §4.1.1 above). In addition to being based on the intuitions of the individual analyst (an issue which can be partly overcome via inter-analyst agreement,

see, e.g., Bybee & Eddington 2006, Zeschel 2012), the main drawback of manual annotation is that it does not produce precise quantitative data that allow for the estimation of degrees of coherence (variability) or similarity. A more systematic (although still not truly quantitative) approach to the classification of verbs is found in Levin (1993). She argues that verbs that are similar in meaning show similar syntactic behaviour. Verbs can be categorised in classes on the basis of the diathesis alternations they do or do not participate in (e.g. the causative/inchoative alternation, conative alternation, dative alternation, locative alternation...), and the meaning aspects these alternations are sensitive to (e.g. contact, motion, change of state...). Although the verb classes are defined for English and based on English alternations, Levin proposes that they may transfer onto other languages as well; the specific verbs and alternations may be different, but they are often sensitive to the same verbal aspects and share the same meaning components (Levin 1993: 10). The verb classes defined by Levin have been very influential in later works and are at the basis of VerbNet, the largest online verb lexicon currently available for English (Kipper-Schuler 2006). Still, it needs to be pointed out that Levin's classification is based on illustrative examples and existing work on said alternations. She admits that "equally valid classification schemes might have been identified instead of the scheme presented in Part II of the book" and that her book is "by no means a definitive and exhaustive classification of the verb inventory of English" (Levin 1993: 18).

More recently, several methods have been developed to establish verb classes in a bottom-up manner on the basis of large-scale corpora. The basic assumption underlying these methods is the distributional semantics hypothesis that words that occur in similar contexts are close in meaning (cf. the well-known phrase "You shall know a word by the company it keeps" by Firth 1957: 11). Most of the analyses consist of two general steps, (i) the identification of contextual clues and (ii) the analysis and interpretation of that contextual information, but there is some difference in the way these steps are performed. In collocational analyses, the analyst uses statistical measures to identify the meaningful collocates in a word's distribution (see, e.g., the association ratio in Church & Hanks 1990). The probability of co-occurrence also lies at the basis of the covarying collexeme analysis, which, as mentioned in §4.2.1, measures the association between two words within the same construction (Gries & Stefanowitsch 2004a, Stefanowitsch & Gries 2005). However, once the significant collocates are identified, the subsequent analysis and interpretation of the semantic patterns are generally still performed manually, so the collocational methods are not especially interesting for "objectively" quantifying the notion of semantic coherence in the productivity model. In contrast, the so-called Behavioural Profile models manually collect distributional information of a certain word (in a specific corpus), but then subject this information to multivariate statistical techniques in order to automatically cluster different senses or establish classes. For instance, Gries (2006) first constructed a corpus-based behavioural profile of *to run*, by manually annotating 815 citations of *run* in British and American English for

morphological features like tense, aspect and voice, syntactic properties like transitive versus intransitive use, objectifiable semantic characteristics (animacy, concreteness, etc.) and by extracting its collocates in the same clause. All this contextual information is brought together in a frequency table, which is then submitted to a hierarchical agglomerative cluster analysis to identify classes of senses. The cluster analysis succeeds in distinguishing the cases of literal motion from the cases of abstract motion, as well as producing multiple small, homogeneous clusters. Glynn (2010) constructed a behavioural profile of the verb *to bother*, annotating 650 occurrences in British and American English for a range of formal and semantic features. Instead of a cluster analysis, he used Correspondence Analysis, which constructs a co-occurrence table for the annotated features and then visualises the calculated relative distances in a graph. Based on the association (proximity in the graph) or disassociation (distance in the graph) of the features, the uses of *to bother* can be grouped in semantically coherent classes, associated with specific syntactic behaviour. The two case studies discussed here are concerned with the polysemy of one particular verb, grouping together its senses on the basis of behavioural features, but the model could reasonably be extended to establishing semantic classes on the basis of the behavioural profile of multiple verbs as well. As has been mentioned in §4.2.1, for example, the hierarchical cluster analysis can also be used to identify classes of multiple verbs on the basis of covarying collexemes (Gries & Stefanowitsch 2010).

While it is not our aim to solve the operationalisation issues here, a promising direction for operationalising semantic coherence in the current productivity model, which should be further explored in future research, comes from the domain of Word Space or Semantic Vector Space models. Word Space models can be document-based, in that the context of a word is defined as an entire given document. This is the case in Latent Semantic Analysis [LSA], a mathematical technique that computes the likelihood that words occur together in the same context, in which "context" is defined as an internally coherent piece of text (see Landauer et al. 1998 for a more technical, detailed step-by-step explanation of the computational mechanisms of LSA). The applications of LSA are numerous (see Landauer et al. 2007), but LSA is relevant for the present discussion because it has already been used in linguistics to approximate speaker's judgments of meaning similarity (i.e. semantic coherence) between words. For example, Suttle & Goldberg's (2011) experimental study on the influence of type frequency, variability and similarity on productivity (see Ch2, §2.1.2) quantifies variability and similarity by determining semantic distance between verbs within the same class and between different verb classes via LSA (the verb classes themselves being pre-defined following Levin 1993). Given that document-based models are based on co-occurrence within larger chunks of text, they are especially interesting when studying basic associative relations. Perhaps better suited for present purposes are the word-based models, in which case the co-occurrence window is delineated by a limited number of words to the left or right of

the target word. There is a difference between token-level models, which are used to measure polysemy (see, e.g., a semasiological case study of the Dutch word *monitor* in Heylen et al. 2015), and type-level models, which can be used to establish verb classes. Without going into the details here, these Vector Space models calculate the (dis)similarity of a selection of target words on the basis of their distributional profile. The most successful models generally only take into account the co-occurring *content* words, which of course requires the use of a part-of-speech tagged corpus.[48] In addition to the so-called "bag-of-words" approach, there are also Vector Space models that take into account the syntactic dependency relationships, but as these can only be performed on syntactically annotated corpora, they are less frequently used. Interestingly, the Vector Space models have already been used to study semantic aspects of productivity more directly (i.e. without immediate reference to the correlation with type frequency). More precisely, it has been investigated to what extent changes in the types of verbs that are associated with certain constructions, like the *way*-construction and the *hell*-construction, reveal shifts in productivity (Perek 2016a, b, see §5.3.2 for some more details). In order to find out how exactly we can integrate these Vector Space models into Barðdal's productivity model, an in-depth evaluation of the method will be required. For example, given that the similarities are calculated on a verb-by-verb basis, it needs to be tested to what extent we can actually use the semantic similarity between verb *pairs* to quantify the semantic coherence of *multiple* verbs used with a specific intensifier; some additional steps (e.g. averaging over all pairwise similarity values) may have to be taken. The current paragraph has offered some possible pathways that are worth exploring in future studies, aimed at including semantic or qualitative aspects of productivity in a more systematic fashion.

4.3.2 Synchronic variation

If we want to examine whether the structure of the constructional network is different in Belgian and in Netherlandic Dutch, we first have to consider the potential differences in the degree of productivity of the subschemas that should or should not be included in the constructional network. As some of the frequency-based measures developed by Baayen and colleagues are highly sensitive to overall corpus size and given the substantial difference in size between the Netherlandic Dutch data set and the Belgian Dutch data set, we will be comparing the two data sets for the largest token size they have in common (Plag et al. 1999: 221). Concretely, we have extracted a random sample of 1,042 tokens (i.e.

[48] The SoNaR corpus is part-of-speech tagged but it is only searchable through the online search module OpenSoNaR, which complicates the automated retrieval of collocates. The Delpher corpus is not (yet) enriched with any linguistic annotation whatsoever. (status April 2018)

the size of the Netherlandic Dutch data set) from the Belgian Dutch data set which will be the basis for our productivity comparison with Netherlandic Dutch. This sample set, despite being so much smaller, is highly representative of the entire data set. The 15 intensifiers that were most frequently used in the entire data set of 2,445 instances are still the top 15 intensifiers in this sample set and the ranking has changed only slightly. The top 15 verbs are also almost identical, with the exception of *zweten* 'to sweat' being replaced by *zuipen* 'to booze' and some minor shifts in the ranking.

4.3.2.1 A frequency-based productivity complex

Before having a closer look at the potential national differences in the productivity of lower-level subschemas, we briefly want to compare the overall productivity of the construction in the two national varieties of Dutch. In the Belgian Dutch sample set of 1,042 tokens, we find 74 different intensifier types, 32 of which are hapaxes, and 103 different verb types, among which 55 hapaxes. When we looked at the full data set of 2,445 tokens for Belgian Dutch in §4.3.2, we concluded on the basis of the type-token ratio that Belgian Dutch had a lower degree of variation of intensifiers (96/2,445=0.04) than Netherlandic Dutch (68/1,042=0.07). At the shared sample size of 1,042, the intensifier slot of the intensifying fake reflexive resultative construction turns out to be slightly more varied and productive in Belgian Dutch (V=74, $\mathscr{P}$=0.03 versus V=68, $\mathscr{P}$=0.02 in Netherlandic Dutch), but the verb slot is more productive in Netherlandic Dutch (V=103, $\mathscr{P}$=0.05 in Belgian Dutch versus V=137, $\mathscr{P}$=0.08 in Netherlandic Dutch). In other words, the construction is more easily extended to new verbs in Netherlandic Dutch than in Belgian Dutch (which may explain why it is overall more frequent in Netherlandic Dutch, cf. supra), but speakers of Belgian Dutch appear to be slightly more creative in their use of intensifiers.

In the remainder of this paragraph, we apply the productivity measures to the partially specified subschemas in order to uncover more detailed, lower-level differences in productivity. Table 4.14 gives an overview of the frequency-based measures for the intensifier-specific subschemas in Belgian Dutch.

Table 4.14. Frequency-based productivity measures for the top 15 intensifiers in SoNaR-BE (sample N=1,042)

	N	V	n1	$\mathscr{P}$	Hapax/type
te pletter	196	47	24	0.12	0.51
blauw	122	6	4	0.03	0.67
rot	102	13	7	0.07	0.54
een hoedje	92	3	2	0.02	0.67
dood	78	13	6	0.08	0.46
uit de naad	70	8	5	0.07	0.63
suf	48	22	18	0.38	0.82
de ziel uit het lijf	37	14	8	0.22	0.57
kapot	30	15	9	0.30	0.60
een bult	22	3	0	0.00	0.00
een ongeluk	22	6	5	0.23	0.83
de longen uit het lijf	15	6	3	0.20	0.50
krom	14	6	4	0.29	0.67
de pleuris	13	7	4	0.31	0.57
een aap	13	2	1	0.08	0.50

We saw in §4.3.2.2 that there are a number of top intensifiers that are exclusively used in one of the two national varieties of Dutch. If we turn to the productivity of these nationally exclusive intensifiers, we find that two of these have a (somewhat) productive use, i.e. *een slag in de rondte* 'a punch around' in Netherlandic Dutch and *de ziel uit het lijf* 'the soul out of the body' in Belgian Dutch. On the other hand, there are intensifiers that are barely productive at all, e.g. *wild* 'wild' in Netherlandic Dutch and *een aap* 'a monkey' in Belgian Dutch. Indeed, the latter two intensifiers were found to enter into near-exclusive collocations with one or two verbs in §4.2. In the section on the organisation of the constructional network (see §4.4 infra), we will discuss the implications of these findings for the overall structure of the network in Belgian and Netherlandic Dutch. In addition to these nationally exclusive intensifiers, there are a number of other intensifiers that are only found in either Table 4.10 or Table 4.14. These intensifiers do occur in both Netherlandic and Belgian Dutch, but they are much more frequent in one of the two varieties than in the other. For example, *een bult* 'a hump' only has one occurrence in the Netherlandic Dutch data set and *krom* 'bent' and *de pleuris* 'the pleurisy' only have three each; conversely, *wezenloos* 'vacant' occurs once and *groen en geel* 'green and yellow' twice in the Belgian Dutch sample set. *Het vuur uit de sloffen* 'the fire out of the slippers' is not in Table 4.14 but only because it just falls outside of the top 15: it also has 13 occurrences in Belgian Dutch, occurring with 3 verb types (among which 2 hapaxes). In the category of overlapping intensifiers that are frequent in both national varieties of Dutch, we see that there are several intensifiers which are approximately equally

"unproductive" in Belgian and Netherlandic Dutch. We note that *een hoedje* 'a little hat' and *blauw* 'blue' have rather low scores for realised and potential productivity, whereas *rot* 'rotten', *dood* 'dead' and *uit de naad* 'out of the seam' perform slightly better on one (or both) of the productivity aspects in both national varieties. Nonetheless, there are also some noticeable differences between some of the top intensifiers. The national differences in productivity emerge more prominently if we visualise them in the global productivity graph in Figure 4.17.

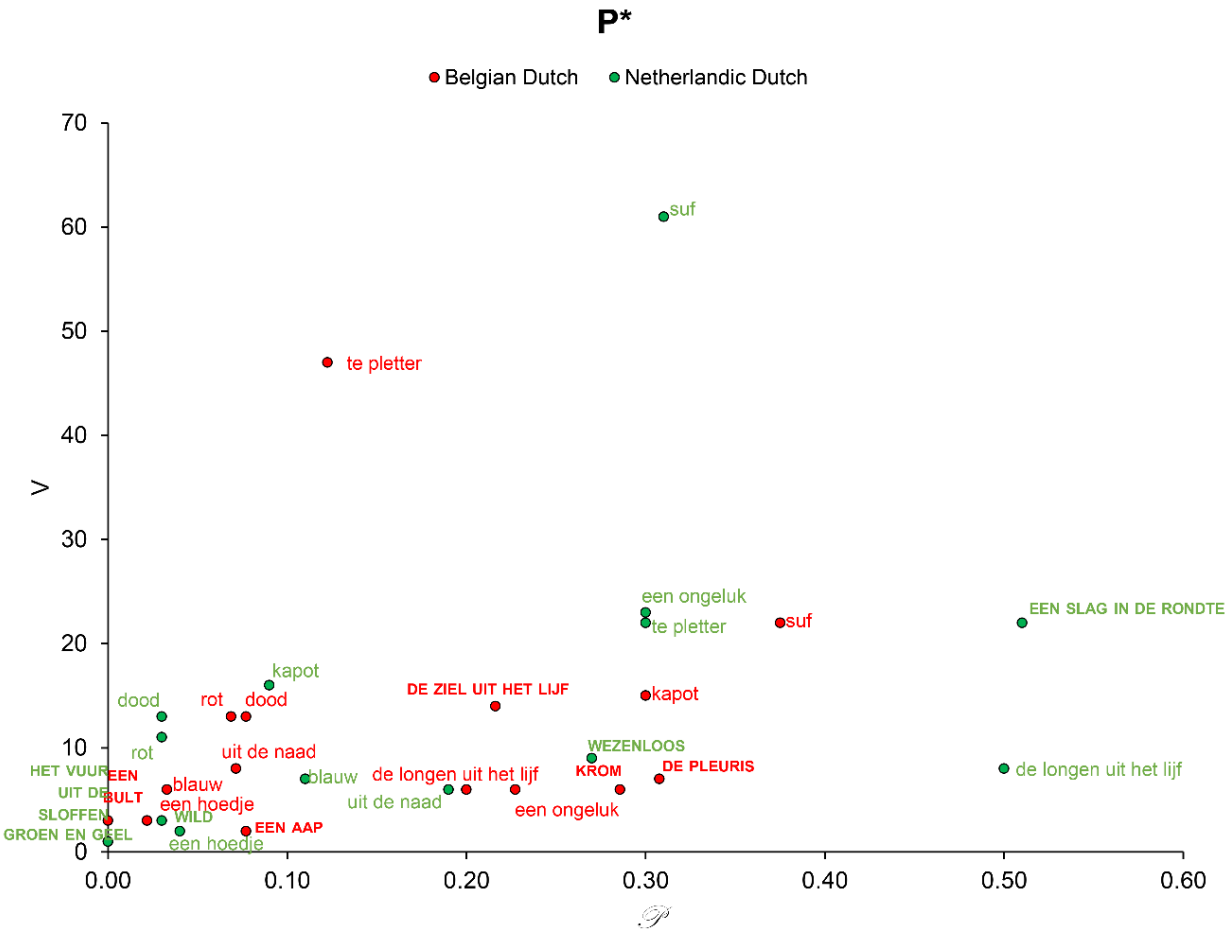


Figure 4.17. Comparison of the global productivity (P*) of the top 15 intensifiers in Belgian and Netherlandic Dutch (N=1,042)

Marked in bold capitals are the intensifiers that occurred in the top 15 of only one of the two national varieties. The graph visually corroborates the results of Table 4.10 and Table 4.14: the intensifiers *groen en geel* (NL), *wild* (NL), *het vuur uit de sloffen* (NL), *een bult* (BE) and *een aap* (BE) are in the bottom left hand corner of the graph, whereas *de pleuris* (BE), *de ziel uit het lijf* (BE), *krom* (BE), *wezenloos* (NL) and *een slag in de rondte* (NL) are positioned much more to the top right of the plane. As for the overlapping intensifiers, the similarities and differences that were discussed earlier are now visually represented by either their proximity or the distance between them. We see that for both national varieties, disregarding some minor fluctuations in their mutual positions, *een hoedje* and

blauw are situated in the bottom left hand corner of the graph and *rot*, *dood* and *uit de naad* are slightly more to the top right. One striking observation is the position of *suf* and *te pletter*: in Belgian Dutch, *te pletter* is the intensifier with the highest V and lower $\mathscr{P}$ and *suf* is characterised by a lower V but higher $\mathscr{P}$, whereas in Netherlandic Dutch, it is more or less the other way around (at least in terms of V). The intensifier *de longen uit het lijf* is situated at approximately the same position on the Y-axis in both national varieties, indicating that it has the same degree of realised productivity. However, the different position on the X-axis indicates that *de longen uit het lijf* is more likely to occur with previously unattested verb types in Netherlandic Dutch. For *kapot*, as well, the vertical position is more or less the same in both national varieties but this time the difference in horizontal position indicates that *kapot* has a higher degree of potential productivity in Belgian Dutch. Finally, *een ongeluk* is situated much more to the top right in Netherlandic Dutch, which suggests that the intensifier is overall more productive in Netherlandic Dutch than in Belgian Dutch. This difference is probably related to the fact that in Belgian Dutch, the collocation with *schrikken* 'to be startled' accounts for 80% of the tokens, whereas it is featured in "only" 40% of the tokens in Netherlandic Dutch.

As before, we should keep in mind that the difference in token frequency between some of these intensifiers may influence the outcome of the productivity measures in important ways (cf. supra). Like in Netherlandic Dutch, the lack of data for most of the individual intensifiers renders it virtually impossible to compare the intensifiers at their common shared sample size, which is extremely low even for the top 15 intensifiers. In order to prove the usefulness of such an approach, we revisit the productivity measures for just the top 5 intensifiers in Belgian Dutch at the same sample size as for Netherlandic Dutch (N = 60) in Table 4.15.

Table 4.15. Frequency-based measures for the top 5 intensifiers at N=60 in SoNaR-BE (sample N=1,042)

	N	V	n1	$\mathscr{P}$	Hapax/type
te pletter	60	23	12	0.20	0.52
blauw	60	4	2	0.03	0.50
rot	60	10	6	0.10	0.60
een hoedje	60	3	2	0.03	0.67
dood	60	11	4	0.07	0.36

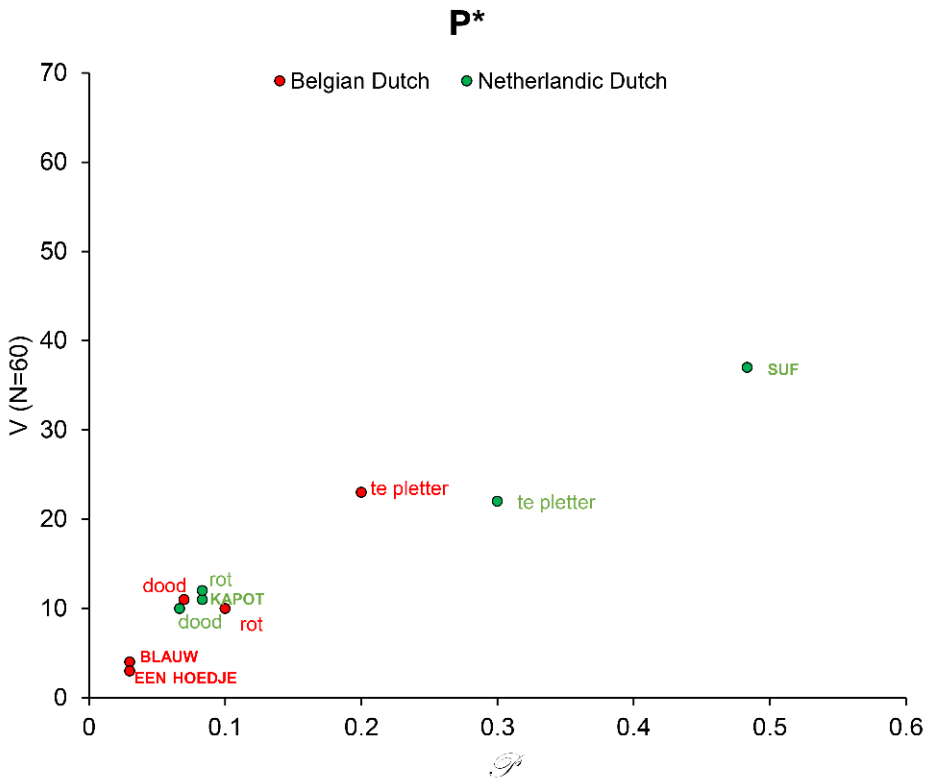


Figure 4.18. Comparison of the global productivity of the top 5 intensifiers at N=60 in Belgian and Netherlandic Dutch (N=1,042)

Compared to Figure 4.17, we find that the Belgian Dutch variant of *te pletter* 'to smithereens' still has a lower degree of realised productivity, but a higher degree of potential productivity at 60 tokens. This suggests that it has not yet saturated its collocational range and still has a higher potential to be extended to previously unattested verb types. Interestingly, the Belgian intensifier *te pletter* is now situated much closer to its Netherlandic counterpart, indicating that at 60 tokens there is not really such a pronounced difference yet. As *te pletter* is much less token frequent in Netherlandic Dutch, with only 60 tokens in total, we do not know whether the Netherlandic variant of *te pletter* would also increase its collocational range if its token frequency was increased. Given its potential productivity, we may expect to see a similar evolution in Netherlandic Dutch if the size of the data set were increased, but it was already pointed out that some caution is warranted in the interpretation of the $\mathscr{P}$-score (see also Ch6, §6.2.2). The other intensifiers in Belgian Dutch have not really reshifted their position much: they already had a relatively low degree of potential productivity at 60 tokens and, to be sure, they have not gained a lot of new types as their token frequencies increased.

We also applied the frequency-based productivity measures to the verbs in Table 4.16.

Table 4.16. Frequency-based productivity measures for the top 15 verbs in SoNaR-BE (sample N=1,042)

	N	V	n1	$\mathscr{P}$	Hapax/type
schrikken	196	17	8	0.04	0.47
werken	148	23	12	0.08	0.52
zich ergeren	105	11	5	0.05	0.45
zich amuseren	75	5	2	0.03	0.40
betalen	68	4	3	0.04	0.75
lopen	58	18	10	0.17	0.56
lachen	57	17	7	0.12	0.41
rijden	30	10	5	0.17	0.50
zich vervelen	28	7	3	0.11	0.43
piekeren	24	5	3	0.13	0.60
zich schamen	19	5	2	0.11	0.40
zoeken	18	5	2	0.11	0.40
drinken	13	6	3	0.23	0.50
schreeuwen	13	6	4	0.31	0.67
zuipen	13	4	2	0.15	0.50

Most of the verbs that are featured in the top 15 of most frequently used verbs in Netherlandic Dutch are also in the top 15 in Belgian Dutch. Three verbs are missing, viz. *eten* 'to eat', *trainen* 'to train' and *rennen* 'to run', which are replaced by *zich amuseren* 'to enjoy oneself', *rijden* 'to ride' and *schreeuwen* 'to yell'. With the exception of *trainen*, which is actually equally token frequent in Belgian and Netherlandic Dutch, the other verbs are indeed remarkably less frequent – although not completely absent – in one of the two national varieties. In §4.1.2.1, it was proposed that some of these frequency discrepancies may be explained by more general differences in the vocabulary of Belgian and Netherlandic Dutch or the type of articles in Belgian and Netherlandic newspapers. Especially noteworthy is the substantial difference for *zich amuseren* 'to enjoy oneself' with 75 occurrences in Belgian Dutch versus only 5 occurrences in Netherlandic Dutch. If we now look at its productivity, we find that *zich amuseren* 'to enjoy oneself' may be very token frequent in Belgian Dutch, but it does not present itself as particularly productive: it occurs with only 5 types overall and its distribution is highly dominated by the two intensifiers *te pletter* 'to smithereens' and *rot* 'rotten', which together account for 70 of the 75 tokens. The verb *rennen* 'to run' does not appear to be very productive either, but given its low token frequency, we cannot really go into much detail here. With respect to the verbs that are frequent in both national varieties of Dutch, the global productivity graph in Figure 4.19 once more facilitates the comparison between Belgian and Netherlandic Dutch.

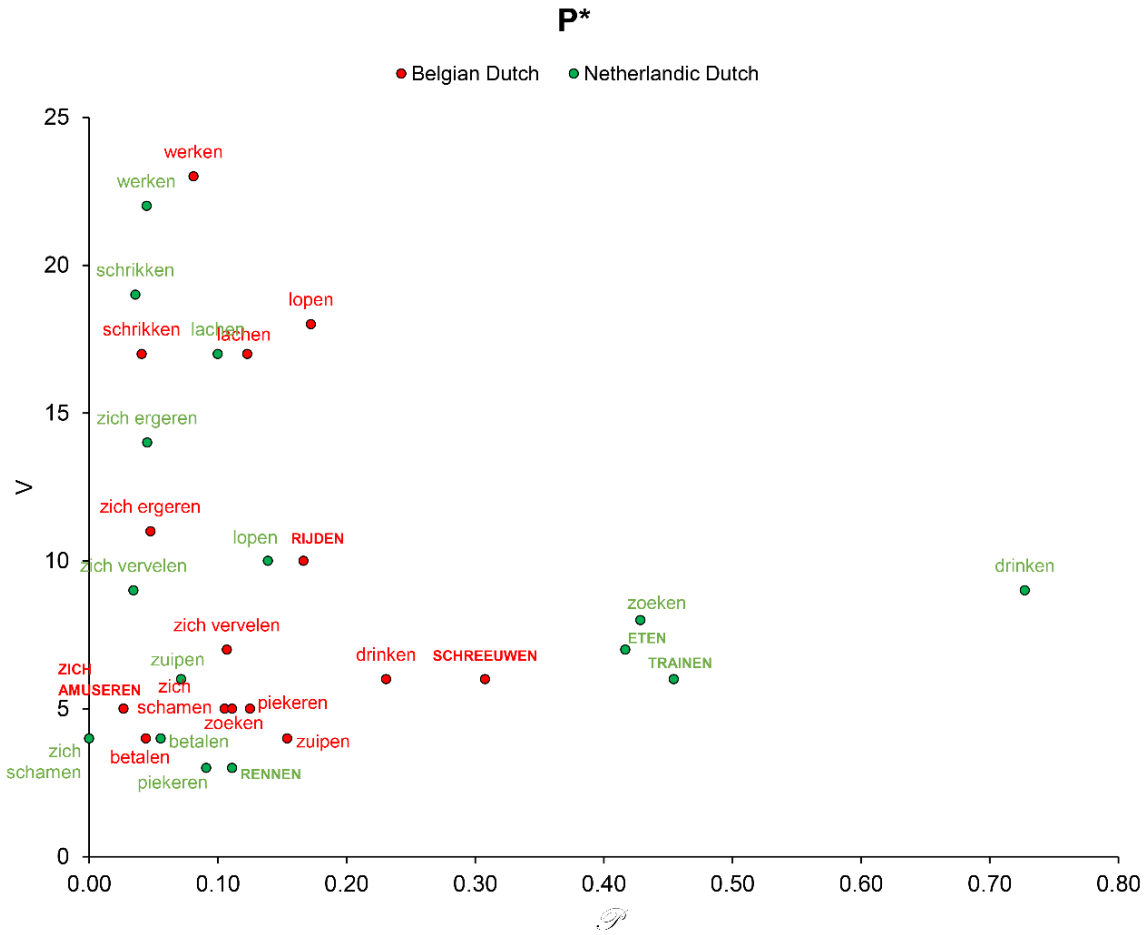


Figure 4.19. Comparison of the global productivity (P*) of the top 15 verbs in Belgian and Netherlandic Dutch (N=1,042)

The overall primarily vertically oriented distribution of the verbs is quite similar in both national varieties of Dutch and several of the national variants of the top verbs are positioned close to one another. In the bottom left area of "low" global productivity, the national equivalents of *betalen* 'to pay' almost coincide and the national equivalents of *piekeren* 'to worry' and *zuipen* 'to booze' are also situated in each other's proximity. Moving up a bit higher on the Y-axis, the Belgian and Netherlandic equivalents of *zich ergeren* 'to be annoyed' and *zich vervelen* 'to be bored' are found within a relatively close distance of one another as well, although both verbs have a lower realised productivity and a (slightly) higher potential productivity in Belgian Dutch. All the way at the top of the Y-axis then, we see that *werken* 'to work', *schrikken* 'to be startled' and *lachen* 'to laugh' are also clustered together at approximately the same coordinates. Focusing our attention on the differences, we can point out that *drinken* 'to drink' and *zoeken* 'to search' are situated more to the top right in Netherlandic Dutch. At approximately the same token frequency, these verbs occur with a wider variety of intensifier types and a higher proportion of hapaxes in Netherlandic Dutch than in Belgian Dutch. The reverse phenomenon is found for the verb *lopen* 'to run', which is globally more productive in

Belgian Dutch as compared to Netherlandic Dutch. The verb *zich schamen* 'to be embarrassed', which had a $\mathscr{P}$-score of zero in the Netherlandic data, is overall much less frequent in Belgian Dutch but it does appear to have at least some degree of potential productivity. For the sake of comparison, we also recalculated the measures for the top 5 verbs in Belgian and Netherlandic Dutch at the common sample size of N=70.

Table 4.17. Frequency-based productivity measures for the top 5 verbs at N=70 in SoNaR-BE (sample N=1,042)

	N	V	n1	$\mathscr{P}$	Hapax/type
schrikken	70	13	6	0.09	0.46
werken	70	16	7	0.10	0.44
zich ergeren	70	10	6	0.09	0.60
zich amuseren	70	5	2	0.03	0.40
betalen	68	4	3	0.04	0.75

The global productivity graph in Figure 4.20 does not reveal any noteworthy differences. Within Belgian Dutch (the red verbs) the relative position of the verbs stays largely the same and if we compare Belgian and Netherlandic Dutch, the top verbs are only slightly closer together at 70 tokens than they were at their total token frequencies.

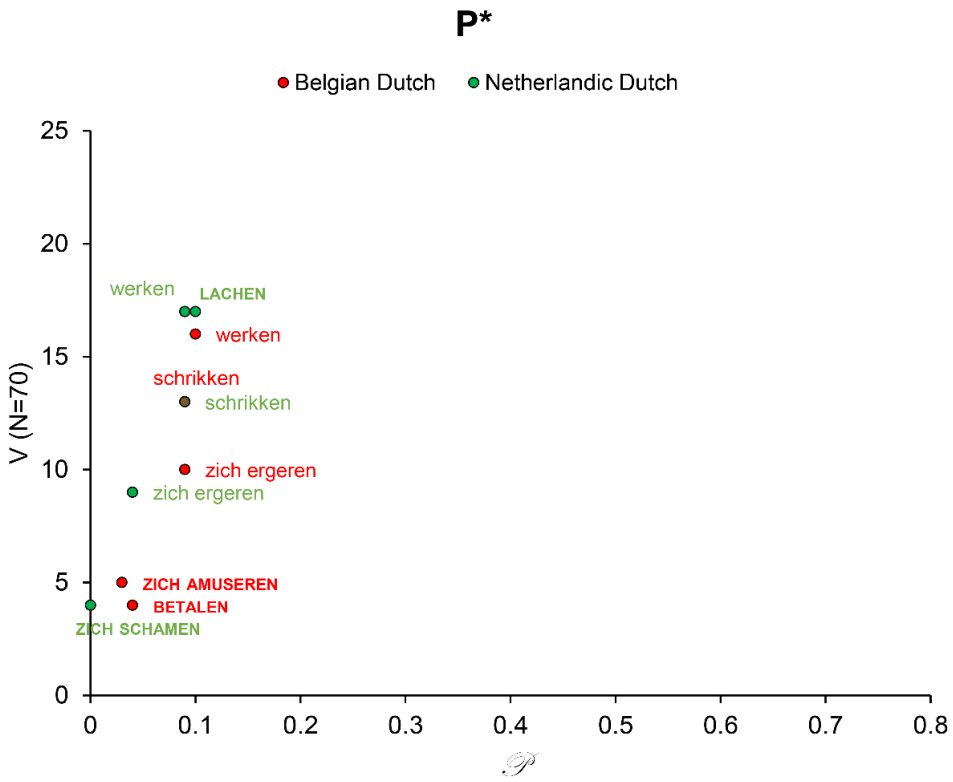


Figure 4.20. Comparison of the global productivity (P*) of the top 5 verbs at N=70 in Belgian and Netherlandic Dutch (N=1,042)

Overall, it would seem that we can distinguish more or less the same four frequency-based productivity classes in both national varieties of Dutch, viz. (i) high V, high $\mathscr{P}$ (ii) high V, low $\mathscr{P}$ (iii) low V, low $\mathscr{P}$ (iii) low V, high $\mathscr{P}$. However, the exact constitution of specific groups may be somewhat different. Several items are of course exclusive to either Belgian or Netherlandic Dutch, but even some of the overlapping members are part of different groups. For example, whereas *de longen uit het lijf* 'the lungs out of body' still shows a lot of growth potential in Netherlandic Dutch (low V, high $\mathscr{P}$), it appears to be much closer to exhausting its productive potential in Belgian Dutch (relatively low V and $\mathscr{P}$). For *kapot* 'broken', the Belgian Dutch variant still shows some growth potential (moderate $\mathscr{P}$), whereas the Netherlandic Dutch variant has a rather low $\mathscr{P}$-score at approximately the same type frequency (V). Both national varieties also have some clear outliers on the Y-axis, viz. *te pletter* 'to smithereens' in Belgian Dutch and *suf* 'drowsy' in Netherlandic Dutch. In the respective national varieties, these two intensifiers are by far the most flexible of all, occurring with a wide array of both frequent and infrequent verb types. It almost appears as if *te pletter* 'to smithereens' in Belgian Dutch and *suf* 'drowsy' in Netherlandic Dutch are the "default" intensifiers (in journalese): if a speaker (or journalist) is unsure which intensifier to use in combination with a new verb, they can hardly go wrong by opting for these respective intensifiers.

4.3.2.2 A constructional model of productivity

Before we can consider the implications of the productivity differences on the overall structure of the network(s), we have to complete the productivity picture by adding information about the semantic coherence of the verbs and intensifiers and the semantic constraints that these items may be subject to.

Again, the semantic constraints are less prominent if we take the perspective of the verb, but there is one interesting case that we did not really notice before. In the Netherlandic Dutch data set, *zich ergeren* 'to be annoyed' occurred with two colour intensifiers *groen en geel* 'green and yellow' and *blauw* 'blue', but the association did not strike us as unusual at the time. However, if we look at the Belgian Dutch data set, we see that *zich ergeren* does appear to have some kind of non-trivial preference for colour intensifiers that other verbs are lacking (cf. §4.4.2 infra). Next to *groen en geel* 'green and yellow' and *blauw* 'blue', we also find *paars* 'purple', *groen* 'green', *spinaziegroen* 'spinach-green', *bicblauw* 'pen-blue' and *donkerblauw* 'dark blue'.

(179) Zelfs wie **zich** soms **bicblauw ergert** aan Jan Fabre, als hij van zijn theater maakt, kan niet ontkennen dat de naald met een vlieg op alles heeft om het hoog te brengen. (SoNaR-BE)
even who himself sometimes pen-blue annoys [...]
'Even those who are sometimes very annoyed by Jan Fabre when he's making a fuss, cannot deny that the needle-with-fly has everything it needs to make it far.'

Taking the perspective of the intensifiers, we find that in Belgian Dutch as well, there are all-round intensifiers that occur with a range of verbs from multiple semantic domains, on the one hand, and more semantically restricted intensifiers that impose certain distributional constraints on the other. What is more, several of the overlapping intensifiers show virtually the same behaviour in both national varieties. The intensifiers *rot* 'rotten', *dood* 'dead' and *kapot* 'broken' belong to the all-round category in both Belgian and Netherlandic Dutch, and they are very similar in type frequency in both national varieties as well. At the other end of the spectrum, *het vuur uit de sloffen* 'the fire out of the slippers' and *de benen uit het lijf* 'the legs out of the body' are virtually limited to the verbs *lopen* and *rennen* 'to run'. Still, we do find one example of *het vuur uit de sloffen* 'the fire out of the slippers' with the verb *trainen* 'to train' in the Belgian data. The context tells us that the subject is talking about training for a cycling competition rather than a running competition, so this appears to be an extension from the original lexical constraint that was found in Netherlandic Dutch.

(180) In de paasvakantie **trainde** ik me echt **het vuur uit de sloffen**, grijnst de student boekhouden-informatica. (SoNaR-BE)
[...] trained I myself really the fire out of the slippers
'In the Easter holidays I really trained my butt off, the student accounting-computer science grins.'

The intensifier *de longen uit het lijf* 'the lungs out of the body' is also productive within a delineated semantic area in Belgian Dutch. Just like in Netherlandic Dutch, it is used with a set of verbs of noise emission or air expulsion and with verbs denoting heavy physical exercise.[49]

(181) Ondergetekende betrad toen met enige schroom De Kuip in Rotterdam, waar die avond ene David Bowie **zich de longen uit het lijf schreeuwde**. (SoNaR-BE)
[...] david bowie himself the lungs out of the body screamed
'With some diffidence, yours truly entered De Kuip in Rotterdam, where one David Bowie screamed his heart out that night.'

[49] Full list of verbs with *de longen uit het lijf* 'the lungs out of the body' in the subset (N=1,042) of Belgian Dutch (for English translations, see the translation list at the beginning of the thesis):
Noise emission/air expulsion verbs: *blazen, schreeuwen, zingen*
Physical exercise verbs: *fietsen, lopen, trainen*

(182) Terwijl de kleine kampioenen **zich de longen uit het lijfje fietsten**, keken de (groot)ouders vertederd toe. (SoNaR-BE)
while the little champions themselves the lungs out of the little body cycled [...]
'While the little champions were cycling their butts off, the (grand)parents were watching with a tender look.'

Uit de naad 'out of the seam', the use of which was found to be similar to *de longen uit het lijf* 'the lungs out of the body', also occurs with physical effort verbs in Belgian Dutch. Typical noise emission verbs are absent, but we do find the verb *repeteren* 'to rehearse' which in the context of (183) refers to singing (note that the one example with *zingen* 'to sing' in Netherlandic Dutch was special as well, cf. supra).[50]

(183) De leden van het koor Canta Libre **repeteren zich uit de naad** om van de koormusical Tafeltje Dek Je later deze maand een spetterende vertoning te maken. (SoNaR-BE)
[...] rehearse themselves out of the seam [...]
'The members of the choir Canta Libre are rehearsing intensely to turn the choir musical "Tafeltje Dek Je" later this month into a smashing performance.'

So far, the interaction between type frequency and semantic coherence appears to lead to similar productivity patterns in both national varieties. Again, however, there are some specific intensifiers that resist straightforward positioning on the productivity cline in Belgian Dutch. For example, *suf* 'drowsy' and *te pletter* 'to smithereens' were found with very different type frequencies in Belgian Dutch and in Netherlandic Dutch, with *suf* 'drowsy' being about three times more type frequent in Netherlandic Dutch and *te pletter* 'to smithereens' being more than twice as type frequent in Belgian Dutch. Still, even in the national variety with the lower type frequency, those intensifiers can still occur with a semantically diverse set of verbs and are thus intuitively rather productive (i.e. 22 types for *suf* 'drowsy' in Belgian Dutch and 22 types for *te pletter* 'to smithereens' in Netherlandic Dutch). If we want to add both national variants of the intensifiers to the productivity cline suggested by Barðdal (2008) (see Figure 4.9, supra), they should all be around the same position on the X-axis (as their verb types barely show any semantic coherence), but suf_{BE} should be much lower on the Y-axis than suf_{NL} and $te\ pletter_{BE}$ should be much higher on the Y-axis than $te\ pletter_{NL}$. In the original model, the large discrepancy in type frequency would make it impossible to put all of these intensifiers *on* the productivity cline, but if we relax the strict linearity of the cline, we can still consider both intensifiers in both national varieties to be highly productive. Another curious case is the Belgian

[50] Full list of verbs with *uit de naad* 'out of the seam' in the subset (N=1,042) of Belgian Dutch: *dansen, fietsen, lopen, repeteren, rijden, spelen* (play a sports*), werken, zwoegen* (for English translations, see the translation list at the beginning of the thesis).

exclusive intensifier *de ziel uit het lijf* 'the soul of the body', which occurs with 14 different verb types – which is one type more than *dood* 'dead' and *rot* 'rotten' – but all of the verbs belong to two semantic classes, viz. the verbs of noise emission, e.g. (184), and the verbs of physical activity, e.g. (185).[51] In that regard, it behaves like *de longen uit het lijf* 'the lungs out of the body', which makes sense in as far as *de longen uit het lijf* 'the lungs out of the body' and *de ziel uit het lijf* 'the soul out of the body' embody a similar imagery of something being drawn or expulsed out of the body. The verb *zweten* 'to sweat' in (186) is not truly a physical activity verb but it is related to that category because it denotes the result of an intense physical activity.

(184) Een groep van negen vrienden en vriendinnen **zingt en orkestreert zich** een weekend lang **de ziel uit het lijf** voor het goede doel. (SoNaR-BE)
[...] sings and orchestrates itself a weekend long the soul out of the body [...]
'A group of nine friends sings and orchestrates their lungs out all weekend for a good cause.'

(185) Iets meer dan vierduizend wandelaars en lopers hebben **zich** vrijdagnacht **de ziel uit het lijf gesport** tijdens de 27ste Nacht van Vlaanderen. (SoNaR-BE)
[...] have themselves friday night the soul out of the body sported [...]
'A little over four thousand hikers and runners have sported their hearts out on Friday night during the 27th Night of Flanders.'

(186) Terwijl zij **zich de ziel uit het lijf zweten**, checken anderen backstage in alle koelte hun mail. (SoNaR-BE)
while they themselves the soul out of the body sweat [...]
'While they are sweating like crazy, the others are backstage in the cool checking their mail.'

Although *de ziel uit het lijf* 'the soul out of the body' in itself appears to be an intensifier that can be easily positioned somewhere in the bottom half of the cline (moderate type frequency and high coherence), its relative position is problematic when compared to other intensifiers, such as *een ongeluk* 'an accident' and *de pleuris* 'the pleurisy', which are found with respectively 6 and 7 semantically *unrelated* types (low coherence) in the Belgian sample set. Although *een ongeluk* 'an accident' is much less type frequent than in Netherlandic Dutch and its distribution in Belgian Dutch is heavily dominated by the verb *schrikken* 'to be startled', it still occurs with 5 other verbs from different semantic categories, viz. *consumeren* 'to consume' (consumption), *kakelen* 'to cackle' (noise emission), *lachen* 'to laugh' (experience) and *meppen* 'to smack' (activity).

[51] Full list of verbs with *de ziel uit het lijf* 'the soul out of the body' in the subset (N=1,042) of Belgian Dutch (for English translations, see the translation list at the beginning of the thesis):

Noise emission/air expulsion verbs: *gillen, juichen, schreeuwen, spelen* (play an instrument)*, zingen*
Physical exercise verbs: *lopen, orkestreren, ploeteren, rijden, spelen* (play a sports)*, sporten, spurten, werken, zweten*, *zwoegen*

(187) Wij **meppen ons** hier **een ongeluk** om die vliegjes te verjagen. (SoNaR-BE)
we smack ourselves here an accident [...]
'We are smacking around like crazy to chase away the little flies.'

(188) Hij zoende de camera en zichzelf, danste de flamenco met bondscoach Ballerini, **kakelde zich een ongeluk.** (SoNaR-BE)
[...] cackled himself an accident
'He kissed the camera and himself, danced the flamenco with national coach Ballerini and cackled like a chicken.'

De pleuris 'the pleurisy', then, is almost evenly distributed across 7 semantically diverse verb types, viz. the physical activity verbs *lopen* 'to run', *rijden* 'to ride' and *werken* 'to work', the experience verb *schrikken* 'to be startled', the communication verb *telefoneren* 'to call', the noise emission verb *zingen* 'to sing' and the general activity verb *sparen* 'to save'.

(189) Mijn grootoom timmerde een huisje voor mij met een sleuf in het puntdak en ik **spaarde me de pleuris**. (SoNaR-BE)
[...] and I saved myself the pleurisy
'My great-uncle built me a little house with a slot in the pointy roof, and I stashed away all my money in it.'

(190) Het gonsde van de geruchten, ik heb **me de pleuris getelefoneerd**. Op het stadhuis kregen ze er de zenuwen van. (SoNaR-BE)
[...] I have myself the pleurisy phoned [...]
'There were rumours buzzing, so I kept calling around. It gave everyone at city hall the jitters.'

It appears that although *de ziel uit het lijf* 'the soul out of the body' is constrained by its original lexical semantics, the productivity within its delimited semantic domain actually produces a higher type frequency than is the case for the all-round intensifiers *een ongeluk* 'an accident' or *de pleuris* 'the pleurisy', the extensibility of which is not a priori semantically constrained. The original productivity model would have us conclude that *een ongeluk* 'an accident' and *de pleuris* 'the pleurisy' are unproductive intensifiers because they are situated below the cline (low type frequency and low coherence), but the data are much less problematic if we also allow for local islands of productivity further away from the cline.

Another possible example of such a productivity island that is not necessarily motivated by semantic coherence is *blauw* 'blue'. Just like in Netherlandic Dutch, *blauw* 'blue' enters into two collocations with two unrelated verbs *zich ergeren* 'to be annoyed' and *betalen* 'to pay', but is also used with other verbs that are not necessarily related to these two frequent verbs, viz. *communiceren* 'to communicate', *doperen* 'to do drugs', *huren* 'to hire' (related to *betalen* 'to pay') and *werken* 'to work'.

(191) Om indruk te maken op de sportdirecteur, **doperen** de kandidaten **zich blauw** voor de rittenwedstrijden van de maand juni. (SoNaR-BE)
[...] (take) dope the candidates themselves blue [...]
'To impress the sports director, the candidates take tons of dope for the stage races in June.'

(192) Vooral als zelfstandige **werk** je **je blauw** en als je niet genoeg investeringen doet, **betaal** je **je** ook nog **blauw** aan belastingen. (SoNaR-BE)
especially as self-employed work you yourself blue [...] pay you yourself also blue [...]
'Especially the self-employed are working hard, and if you don't invest enough, you also have to pay a lot of taxes.'

In example (192), *zich blauw werken* 'to work oneself blue' is probably used in parallel with the more conventional expression *zich blauw betalen* 'to pay oneself blue' a bit further in the sentence. This shows that, in addition to the influencing factors we have mentioned so far, the context may also play a role in the choice of a particular verb-intensifier combination (see also the creative contextual uses in earlier sections, i.e. (145), (146) and (148) to (150)). Curiously, we also find some very local analogical extensions that are not always easily explained by semantic (or formal) similarity. The intensifier *een hoedje* 'a little hat', which was said to form an almost exclusive collocation with *schrikken* 'to be startled', also has two attestations with other verbs. Although the combination of *een hoedje* 'a little hat' with the verb *lachen* 'to laugh' could be explained by referring to the semantic domain of experience verbs, the same cannot be said for the combination in (193) below.

(193) Marino **tapt zich** zoals gewoonlijk **een hoedje** in de studentenzaak De Fitlink. (SoNaR-BE)
marino taps himself as usual a little hat [...]
'As usual, Marino is tapping beer after beer in the student's bar De Fitlink.'

The combination of *zich een hoedje tappen* 'to tap oneself a little hat' presents a challenge to most of the explanatory factors that we have invoked thus far. *Tappen* is a verbal hapax, so we cannot rely on its token or type frequency here (in contrast to *lachen* 'to laugh', the co-occurrence of which could also be explained by referring to the overall flexibility of the verb, cf. §4.2.2 supra), and it is not semantically related to the other two verbs occurring with *een hoedje.* Such apparently "random" extensions are somewhat unexpected. Even if we were to no longer adopt a strict inverse correlation between type frequency and semantic coherence, it is still the case that novel coinages of a schema are generally hypothesised to reflect the semantic space that is already covered by the schema (e.g. Zeschel 2012: 185). In the next section, we will propose that within a constructional network approach, in which schemas are posited at different levels of schematicity and can be taxonomically linked to multiple higher-order schemas through

inheritance links, the non-conventional examples with *blauw* 'blue' and *een hoedje* 'a little hat' are not as puzzling as they may appear at first sight.

4.4 A multiconfigurational network approach to the intensifying fake reflexive resultative construction in present-day Dutch

The taxonomic structure of the constructional network is the result of schema-formation or schematisation, a process by which language users perceive and generalise over formal or semantic similarities between specific linguistic expressions and capture these in a schema. In Chapter 2, §2.1.3, we saw that micro-constructions are abstractions over specific constructs and that subschemas may arise as generalisations over those micro-constructions. Moving up, subschemas may also abstract over other, lower-level subschemas, before eventually arriving at the most abstract level, viz. that of the schema. The main aim of this section is to demonstrate how the results from the multidimensional productivity analysis in the previous section can guide the linguist in constructing the constructional network of the intensifying fake reflexive resultative construction. We have illustrated that subschemas may differ in their ability to sanction novel, previously unattested instances. Given that productivity is intertwined with the abstractness or schematicity of a pattern, we now aim to show how these subschemas, displaying various degrees of productivity, fit into the overall structure of the constructional network. A number of questions will be answered: which abstractions should we take into consideration based on our corpus data (i.e. at which level in the hierarchy should we posit the different subschemas), how do the apparently unproductive intensifiers or verbs fit into the network, how can we account for seemingly "unmotivated" instances that are not immediately subsumed by a superordinate subschema, how are the different subschemas linked to one another?

It is important that one distinguishes between the constructional network as a linguist's heuristic tool to present a construction in a taxonomic hierarchy (often based on concrete corpus data), on the one hand, and the constructional network as the cognitive representation of all knowledge that a language user has about a specific construction, on the other. While most cognitive linguists are aware of the distinction, it is not always made explicit. Blumenthal-Dramé (2012: 41) even states that "there is an unfortunate blurring between the linguist's theoretical metaknowledge and the speaker's internal language system." One crucial aspect that has not been sufficiently emphasised in the existing works on constructional networks is that, during the process of

constructing a visual representation of the constructional network, the linguist has to make several methodological choices that are not necessarily relevant to the language user. For one, the linguist has to decide which slot to focus on when constructing a taxonomic network. Although in the case of clause-level constructions the (main) verb often presents itself as the obvious candidate (cf. Ch2, §2.1.2 and §2.1.3), the choice is less straightforward in the case of the intensifying fake reflexive resultative construction. We have demonstrated in the previous sections that both the perspective of the verb and the perspective of the intensifier provide valuable information on the use and productivity of the construction, and, in addition, it was found that both slots seem to interact in non-trivial ways. Focusing our attention on just one slot in the constructional network would be ignoring the specific nature of the intensifying fake reflexive resultative construction. Still, when building the constructional network in a usage-based, bottom-up fashion, it seems that the linguist is presented with that choice quite early on. Concretely, we start out at the lowest level of the network with micro-constructions in which both the verb and intensifier are lexically specified. Moving up to the subschema level, we can either generalise over the verb slot and get intensifier-specific subschemas with open verb slots [SUBJ V REFL *specified intensifier*], or we abstract over the intensifier slot and get verb-specific subschemas with an open intensifier slot [SUBJ *specified verb* REFL INT]. Effectively, this is where we, when trying to build a visual, taxonomic representation of the network, need to decide whether we are primarily going to focus on the intensifier slot or on the verb slot, and the path that is chosen will determine the further outline of the network structure. Language users, on the other hand, do not have to choose just one path. Instead, they can make several generalisations at the same time, thus simultaneously gaining access to multiple configurations of the network that are expected to be in constant interaction. Which generalisations are perceived of as relevant may not only differ from one language user to another, based on his/her linguistic experience with the construction, but also from one usage context to another. We could conceive of the cognitive constructional network as one giant, complex network in which all configurations come together and in which the different nodes are connected not only via taxonomic relationships but also other types of links that can only be fully appreciated in a multidimensional network.

In addition, there are certain aspects of use of the construction that are part of the language user's knowledge about the construction – thus intrinsic to the language user's *cognitive* representation of the constructional network – but that do not necessarily have a place in the linguist's *taxonomic* representation of the network. For example, it has been mentioned earlier in this chapter that some intensifiers are felt to be more "expressive" or have a stronger intensifying force than others, which may affect their use in concrete situations. It is not unlikely that different intensifiers are associated with different kinds of discourse contexts: depending on the subject of the article, for instance, some intensifiers may be deemed more or less appropriate than others. Such kinds of context-

related usage effects or connotations are not visualised in the network, but they are of course important in the everyday use of the construction. At any rate, any two-dimensional taxonomic representation of the network will always be an oversimplification of the actual dynamics and complexity of constructional networks in the cognitive space of the language user. It needs to be emphasised that, even though we may not currently be able to *visualise* all these different interrelationships and usage details in our taxonomic hierarchies (although computational models are being developed to account for the dynamic structure of constructional networks cf. van Trijp & Steels 2012), we can still take a theoretical multiconfigurational approach to the constructional network structure. We will show that is possible to explain certain peculiarities or low-level idiosyncrasies in the intensifying fake reflexive resultative construction – and, by extension, other constructions – by referring to interactions between different configurations of the network.

Although the linguist's taxonomic hierarchy is a theoretical abstraction (and simplification) that transcends the individual language user, it is meant to provide some insight into the way in which the construction is represented and organised in the minds of language users in general. As the representation of the network is based on concrete corpus data, this raises the question of to what extent we can really gain insight into the cognitive representation of linguistic units from corpus data (Blumenthal-Dramé 2012). The central topic in this debate concerns the notion of entrenchment, i.e. the strength with which certain linguistic units are represented in the minds of the speaker. The so-called Corpus-To-Cognition-Principle by Schmid (2000: 38-39) states that corpus data offer a gateway to cognitive entrenchment patterns via statistical generalisations over text frequencies. This principle is at the basis of many corpus linguistic methods like the already-mentioned distributional semantics methods and collostructional analyses. At the same time, it has been criticised for being too deterministic and making certain assumptions that are, in themselves, not unproblematic (as has recently been admitted by Schmid 2010 himself). Glossing over the details of the debate, the main problem lies in the assumption that corpus data could provide a perfect one-to-one mapping of the cognitive organisation of language. Even so, there is sufficient empirical evidence to suggest that corpus data - representing real language use, produced by real speakers (i.e. the corpus-as-output view in Stefanowitsch & Flach 2017) – do to some extent mirror the linguistic knowledge of language users (Blumenthal-Dramé 2012: 44-65). We will return to the cognitive reality of our network proposals in Chapter 6, §6.2.3. For now, suffice it to say that we adopt a nuanced version of the Corpus-To-Cognition Principle, and rely on the text frequencies in our corpus data in order to propose careful, tentative entrenchment degrees for the patterns at different levels in the hierarchy. The entrenchment indications are primarily aimed at illustrating the relation between different types of micro-constructions and subschemas and investigating the relationship between entrenchment and productivity (see also Ch6, §6.2.3). In usage-based models, the

degree of entrenchment is traditionally related to text or token frequency, in that the frequent repetition of a specific instance of a pattern may lead to the entrenchment of that very instance (Langacker 1987, Bybee 2003, Goldberg 2006, Lieven & Tomasello 2008, inter alia). Importantly, (partially) abstract or schematic units can also be entrenched, in which case it is the type frequency (i.e. the number of different instantiations) rather than the token frequency (i.e. the frequency of specific instantiations) that is decisive. A detailed account of measuring entrenchment at different levels of complexity and schematicity on the basis of corpus data has recently been proposed by Stefanowitsch & Flach (2017). They suggest that the entrenchment of units depends on their "usage intensity", a term which has a slightly different interpretation for different types of units. They argue that high token frequency alone is not always a good predictor of entrenchment, especially if we move beyond the simplest units of language. Instead, we could rely on association-based measurements as an indicator of the cognitive status of multi-word units (cf. collostruction strength, §4.2): rather than just looking at the raw frequency of co-occurrence of multiple items, one could look at items that co-occur more frequently than expected, thus also incorporating the *salience* of a collocation into the analysis (see also Gries & Stefanowitsch 2004b, Stefanowitsch & Gries 2003, 2005). In addition, they suggest to use the type-token ratio rather than just type frequency as an indication of the entrenchment of schematic units. This has the advantage of also taking into account token frequency – as type frequency is to some extent dependent on token frequency –, but the use of token-based ratios comes with a certain downside as well. The ratio may overestimate the entrenchment of low token frequency schemas or underestimate the entrenchment of extremely high token frequency schemas.

The notion of entrenchment has been directly related to productivity. For example, Clausner & Croft (1997) distinguish three options (based on what they call the Bybee/Langacker account). If a specific combination occurs frequently enough, the token can become entrenched as an autonomous chunk, in which case no superordinate schema is posited and there is no productivity (option 1). It is also possible that a limited number of high frequency tokens still give rise to an overarching abstraction/schema. This results in "low productivity", in which case the specific highly token frequent instances are more entrenched than the schema with low type frequency (option 2). The third option is full productivity: a large number of lightly or non-entrenched tokens are part of a highly entrenched, type frequent schema. Barðdal (2008: 49) partly agrees with Clausner & Croft, but she adds that the difference between different degrees of productivity (or low and high productivity) "is not only a difference in the entrenchment of a superordinate schema, but rather a difference in both the type frequency of the two schemas and in the degrees of schematicity at which they exist". Accordingly, the productivity of a construction is said to be determined by its highest level of schematicity. We will show that, given that the previous sections have revealed substantial variation between individual verbs and intensifiers, it may actually be more informative to look at lower

levels of abstraction when investigating (pockets of) productivity than to focus on the most abstract level of the network. In what follows, we will use a combination of absolute token frequency and collostruction strength as the primary indicator of entrenchment at the level of the micro-construction and type frequency as an indicator of entrenchment at the (sub)schema level.

4.4.1 Synchronic use

In this paragraph, we will build two possible taxonomic network representations of the intensifying fake reflexive resultative construction in Netherlandic Dutch, one in which we opt for low-level intensifier-specific subschemas (henceforth the intensifier-centred network) and one in which we opt for low-level verb-specific subschemas (henceforth the verb-centred network). Other possible representations and potential interactions between multiple configurations will be referred to when relevant. Concretely, we will propose a step-by-step procedure towards building a (partial) representation of the constructional network of the intensifying fake reflexive resultative construction. Starting out with a very basic skeleton of the network structure, each step will add further details on the basis of the data presented in the previous sections. As was mentioned above, we do not have direct access to the language user's cognitive network and corpus data necessarily only provide an abstract and simplified picture of the generalisation strategies that individual language users rely on. For example, we do not know exactly how many different types language users need to have experienced before forming a generalisation in the form of a subschema, nor do we know how fine-grained their distinctions really are (see Ch6, §6.2.3.2 for more discussion). As was mentioned above, it is very likely that there are both inter-subjective and intra-subjective differences with respect to the similarities that are perceived as relevant and the readiness to generalise over specific instances. In the case of the intensifying fake reflexive resultative construction, it is quite likely that the specific verb or intensifier also have some impact on the kinds and number of intermediate generalisations that are made (cf., e.g., the so-called "ad-hoc" constraints for some intensifiers in §4.3.1.2), which causes some subareas of the network to be more densely populated than others. Our aim is to illustrate how frequency data, such as token frequencies, type frequencies and collostruction strength, in concert with the information on semantic coherence, can already give us a general idea of the taxonomic organisation of the constructional network of the intensifying fake reflexive resultative construction.

At the lowest level in the network, we could start out with the constructs, a level at which each of the 1,042 individual attestations of the [SUBJ V REFL INT] pattern would have to be represented. Even if we abstract over the specific subject and reflexive pronoun in the form of a lexically specified micro-construction, we still find 318 different

verb-intensifier combinations. For reasons of feasibility, we have selected a number of cases that will allow us to illustrate the main factors that come into play when constructing a (or multiple) constructional network representation(s), viz. [SUBJ *werken* REFL *dood*], [SUBJ *lopen* REFL *de longen uit het lijf*], [SUBJ *piekeren* REFL *suf*], [SUBJ *ergeren* REFL *groen en geel*] and [SUBJ *schrikken* REFL *een hoedje*]. If we were to simply generalise over these micro-constructions, without looking at the actual data, we would get the intensifier-specific subschemas [SUBJ V REFL *dood*], [SUBJ V REFL *de longen uit het lijf*], [SUBJ V REFL *een hoedje*] and [SUBJ V REFL *suf*], as in Figure 4.21.[52]

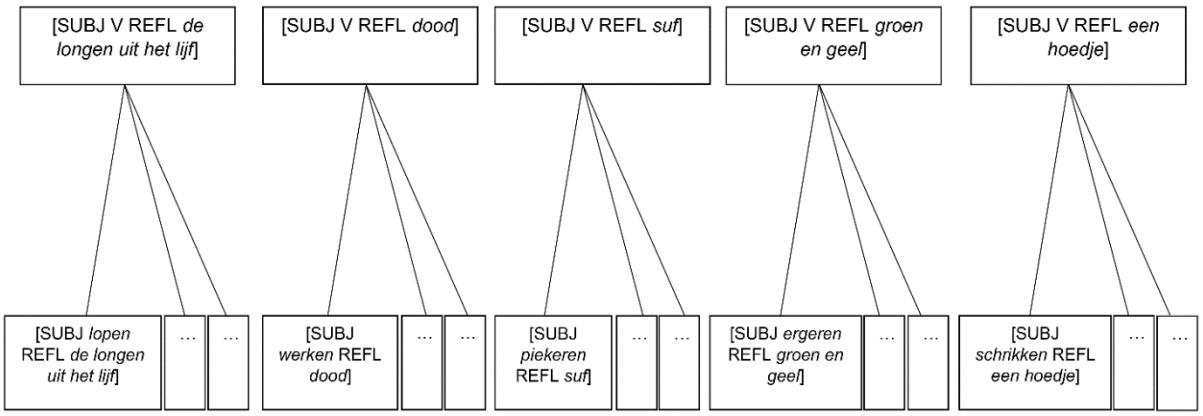


Figure 4.21. Towards building a constructional network, step 1a

Or, we get the verb-specific subschemas [SUBJ *vervelen* REFL INT], [SUBJ *lopen* REFL INT], [SUBJ *schrikken* REFL INT], [SUBJ *ergeren* REFL INT] and [SUBJ *piekeren* REFL INT] in Figure 4.22.

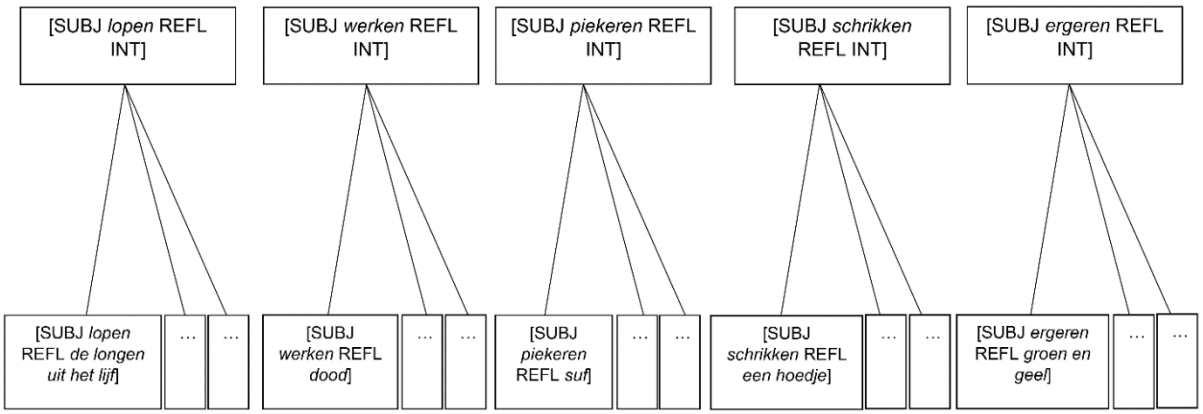


Figure 4.22. Towards building a constructional network, step 1b

Of course, a language user will most likely only form these abstractions if he/she observes that multiple different verbs (or intensifiers) can fill the verb slot (or intensifier slot) of a particular intensifier-specific subschema (or verb-specific subschema). The current

[52] The order in which the micro-constructions and subschemas are presented from left to right is not meant to be meaningful, it is mainly chosen for practical reasons.

representations of the network are still very rudimentary because they assume that all these micro-constructions and subschemas exist at the same level of schematicity and entrenchment. Based on the corpus data presented in the previous sections, however, it is safe to say that this is certainly not the case. First of all, a number of verbs and intensifiers have been shown to be rather productive in that they are able to co-occur with a wide variety of types, e.g. the intensifiers *dood* 'dead' and *suf* 'drowsy' and the verbs *lopen* 'to run', *werken* 'to work', *zich ergeren* 'to be annoyed' and *schrikken* 'to be startled'. In these cases, the subschema is arguably more entrenched than the specific micro-constructions. However, this is clearly not the case for all of these subschemas. In §4.2.1 some intensifiers and verbs were found to be virtually limited to strong collocations: *zich* ***groen en geel*** *ergeren* 'to annoy oneself green and yellow', *zich* ***een hoedje*** *schrikken* 'to startle oneself a little hat' and *zich suf* ***piekeren*** 'to worry oneself drowsy' all had high collostruction strengths. We could state that [SUBJ V REFL *groen en geel*], [SUBJ V REFL *een hoedje*] and [SUBJ *piekeren* REFL INT] can, therefore, not be said to form productive subschemas or that these verbs and intensifiers should not be represented at a subschema level in the hierarchy. Rather, instances like *zich suf piekeren* or *zich een hoedje schrikken* are examples of highly entrenched micro-constructions that have in all likelihood not triggered an overarching abstraction. Barðdal (2008: 48) observes that "constructions can only be assumed to exist at abstract schematic levels if there are linguistic data in support of such an analysis". While she is primarily referring to highly abstract schemas at the macro-level, the idea that we can only posit certain generalisations if we have sufficient linguistic evidence also applies to lower levels. In the context of language acquisition as well, Lieven & Tomasello (2008: 186) have claimed that "[certain] higher-level schemas may only be weakly represented and, indeed, they may sometimes only exist in the formalised grammars of linguists". We adjusted the network representations accordingly by adding a bold frame to entrenched micro-constructions and putting the non-existent subschemas in grey dashed frames in Figure 4.23 and Figure 4.24.

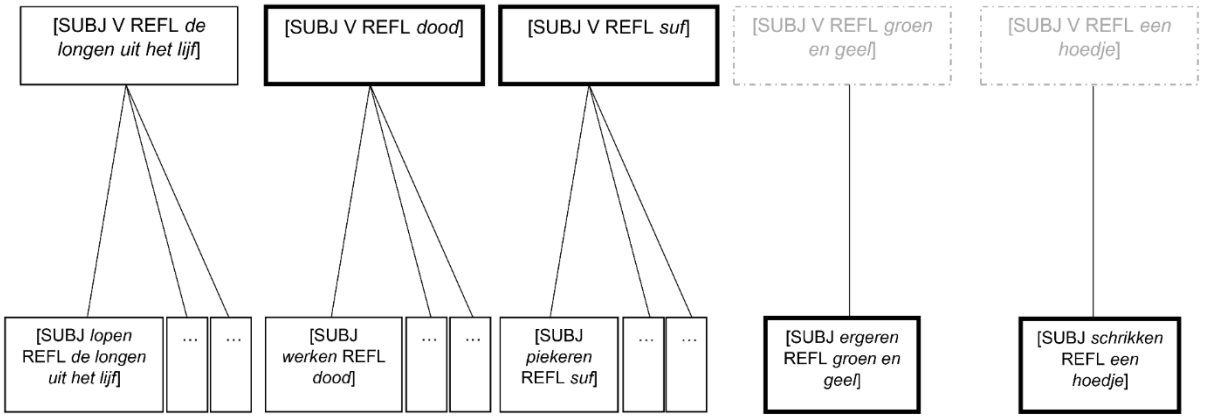


Figure 4.23. Towards building a constructional network, step 2a

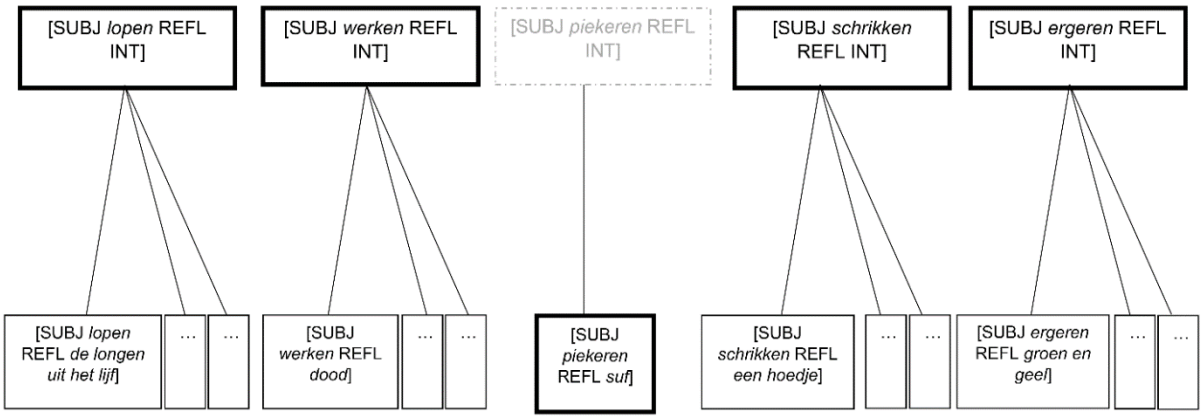


Figure 4.24. Towards building a constructional network, step 2b

Highly token frequent micro-constructions do not necessarily inhibit the creation or productivity of the overarching subschema. Given that a highly entrenched token is in itself also a type, it may be of importance for subschemas at lower levels of schematicity, in as far as it can become available as a model item for analogical extensions, which may in turn give rise to a low-level subschema (cf. Barðdal 2008, Zeschel 2012). There are some indications in our data that certain subschemas that are currently productive might have arisen out of an entrenched micro-construction. For example, we saw that the subschema [SUBJ V REFL *suf*] is very productive in present-day Netherlandic Dutch, but at the same time, there is a set of mental activity verbs (e.g. *denken* 'to think', *dromen* 'to dream', *peinzen* 'to ponder', *rekenen* 'to calculate' and *piekeren* 'to worry' as the most frequent one) that stand out as a rather coherent group among the overall semantically diverse set of verbs. Based on these findings, it may well be possible that the productive subschema [SUBJ V REFL *suf*] has arisen out of a more restricted subschema [SUBJ $V_{\text{mental activity}}$ REFL *suf*], which in turn might have arisen out of the entrenched collocation *zich suf piekeren* 'to worry oneself drowsy'. The diachronic analysis in Chapter 5 will have to explicate whether this is indeed the path that *suf* 'drowsy' has followed over time. At any rate, in the synchronic representation of the network, we could account for the fact that several of the verbs attested in combination with *suf* 'drowsy' belong to a semantically coherent group of mental verbs in the form of an intermediary semantically specified subschema, see Figure 4.25.

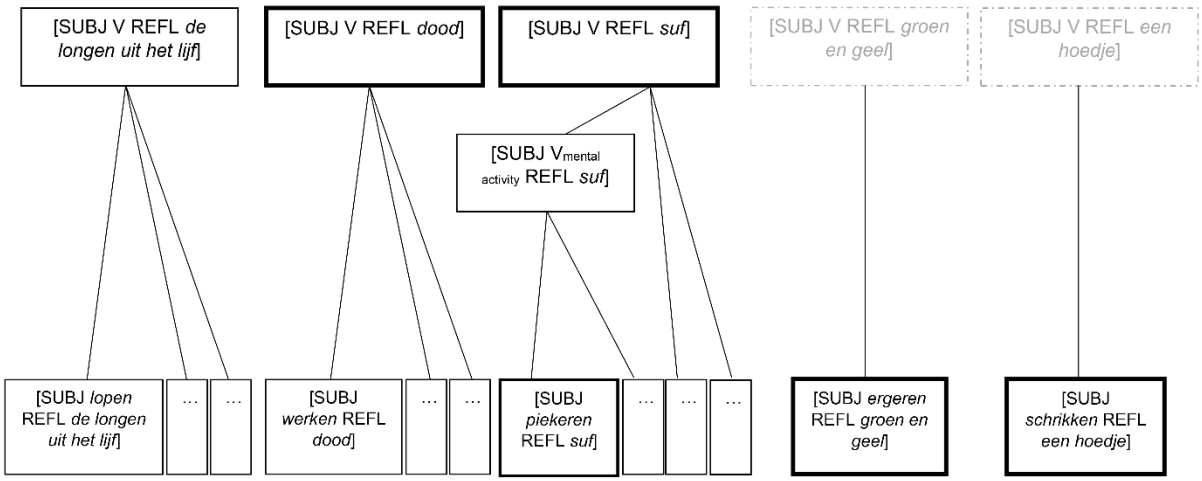


Figure 4.25. Towards building a constructional network, step 3a

The benefit of assuming multiple interacting representations of the network is demonstrated by cases like [SUBJ *schrikken* REFL *een hoedje*] and [SUBJ *ergeren* REFL *groen en geel*], which, for now, appear to be "isolated" micro-constructions in the intensifier-centred representation of the network, but are perfectly motivated in the verb-centred network in Figure 4.26. That is, they are motivated by the subschemas [SUBJ *schrikken* REFL INT] and [SUBJ *ergeren* REFL INT], respectively. Even though they are now licensed by a productive, type frequent (i.e. entrenched) subschema, these micro-constructions are still strong collocations that need to be represented as entrenched in the verb-centred network as well. Similarly, in the verb-centred network, the limited combinatorial flexibility of *piekeren* does not result in the verb-specific subschema [SUBJ *piekeren* REFL INT]: the micro-construction [SUBJ *piekeren* REFL *suf*] appears to be an island in the verb-centred network, even though it was directly licensed by [SUBJ $V_{\text{mental activity}}$ REFL *suf*] in the intensifier-centred network in Figure 4.25.

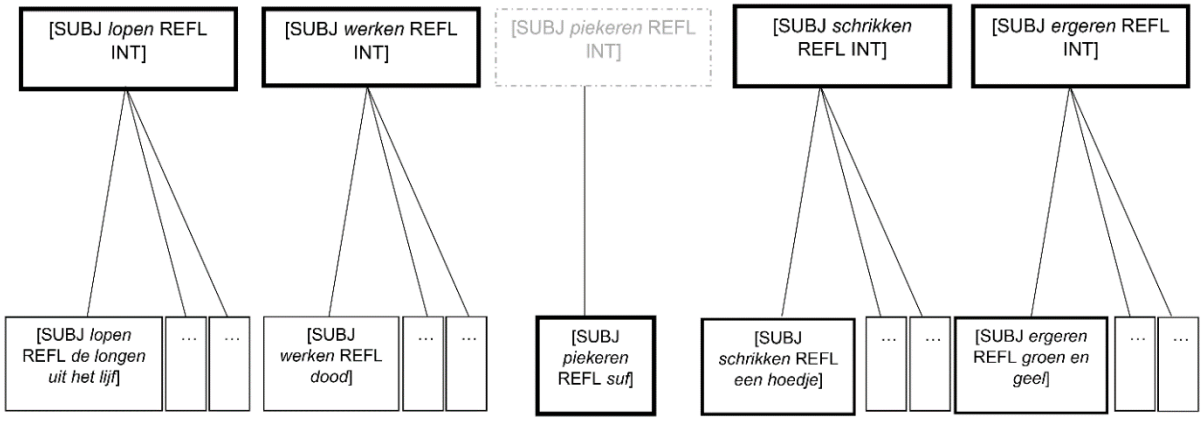


Figure 4.26. Towards building a constructional network, step 3b

If anything, these results suggest that even though a micro-construction may be entrenched in the sense that it is accessible as a whole, it can still be analysed as a compositional combination of an individual verb and an intensifier, i.e. as an instance of

a higher-order (sub)schema in at least one representation of the network. We will return to the notion of entrenchment and cognitive representation in more detail in Chapter 6, §6.2.3. These kinds of network interactions can also explain why the verbs and intensifiers that nearly exclusively occur in such entrenched micro-constructions (i.e. *zich* ***een hoedje*** *schrikken* 'to startle oneself a little hat' or *zich suf* ***piekeren*** 'to worry oneself drowsy') can still display slight variation, as demonstrated by examples (194) and (195).

(194) De sterfscène [...] is om **je een hoedje** te **lachen**. (SoNaR-NL)
the dying-scene is to yourself a little hat to laugh
'The dying scene is so hilarious.'

(195) Daarover **pieker** ik **me rot**. Puzzelend, wikkend en wegend bereidt Advocaat zich deze week voor. (SoNaR-NL)
about that worry I myself rotten [...]
'I worry about it a lot. While he sits puzzling and pondering, Advocaat is preparing himself this week.'

These variations are at first sight somewhat puzzling: in the absence of the subschemas [SUBJ V REFL *een hoedje*] and [SUBJ *piekeren* REFL INT], we might not have expected *een hoedje* 'a little hat' or *piekeren* 'to worry' to show up with any other items outside of the entrenched micro-constructions. In a multiconfigurational model of networks, a possible explanation presents itself. Just like the entrenched micro-construction *zich een hoedje schrikken* 'to startle oneself a little hat' is not an instance of a subschema [SUBJ V REFL *een hoedje*] (which does not "exist", cf. supra) but of the subschema [SUBJ *schrikken* REFL INT], so is *zich een hoedje lachen* 'to laugh oneself a little hat' a micro-construction of [SUBJ *lachen* REFL INT]. In the same vein, where *zich suf piekeren* 'to worry oneself drowsy' is licensed by [SUBJ $V_{\text{mental activity}}$ REFL *suf*] instead of [SUBJ *piekeren* REFL INT] (which was also assumed not to exist, cf. supra), *zich rot piekeren* 'to worry oneself rotten' is motivated by [SUBJ V REFL *rot*] in another possible representation of the network, which prioritises other generalisations. In other words, we can posit certain interactions between the intensifier-centred and verb-centred networks for the "unconventional" micro-constructions *zich een hoedje lachen* 'to laugh oneself a little hat' and *zich rot piekeren* 'to worry oneself rotten' just like we did for the conventional, entrenched micro-constructions *zich een hoedje schrikken* 'to startle oneself a little hat' and *zich suf piekeren* 'to worry oneself drowsy'. This is visualised by the dashed red lines in Figure 4.27.

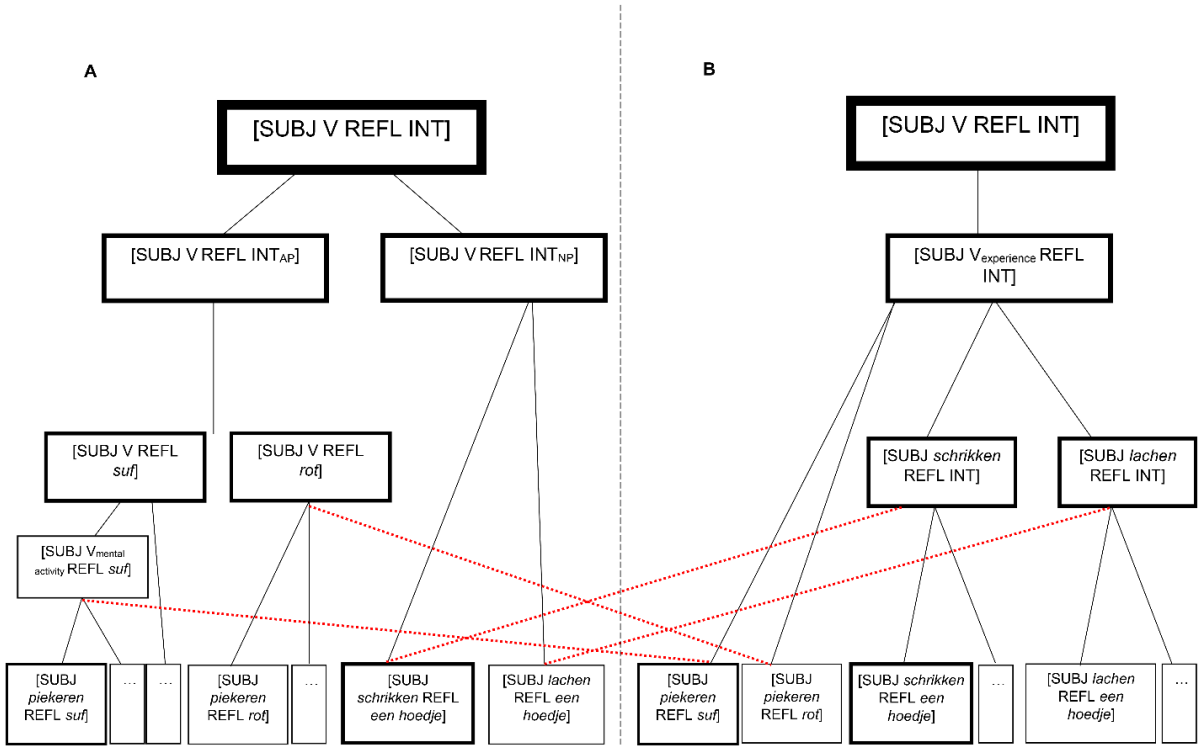


Figure 4.27. Network interactions between intensifier-centred network (A) and verb-centred network (B)

If we continue to build our network, we still need to address the representation of *de longen uit het lijf* 'the lungs out of the body' in the intensifier-centred network. In the previous section, it was observed that this intensifier was semantically constrained in that it only co-occurred with two specific sets of verbs, viz. a semantically coherent group of verbs of noise emission and a coherent group of verbs of physical exercise. This indicates that the subschema [SUBJ V REFL *de longen uit het lijf*] is too general, because the verb slot is not actually completely "open". If we want to remain true to our data, we at least need two lower-level subschemas [SUBJ V$_{noise\ emission}$ REFL *de longen uit het lijf*] and [SUBJ V$_{physical\ exercise}$ REFL *de longen uit het lijf*]. In as far as the language user perceives the shared element of lung capacity (which was admittedly posited as a rather ad-hoc constraint in §4.3.1.2 above), we could also add the [SUBJ V$_{lung\ capacity}$ REFL *de longen uit het lijf*] as an overarching schema, but we opted for a dashed frame in Figure 4.28 to indicate the uncertain status of this subschema.

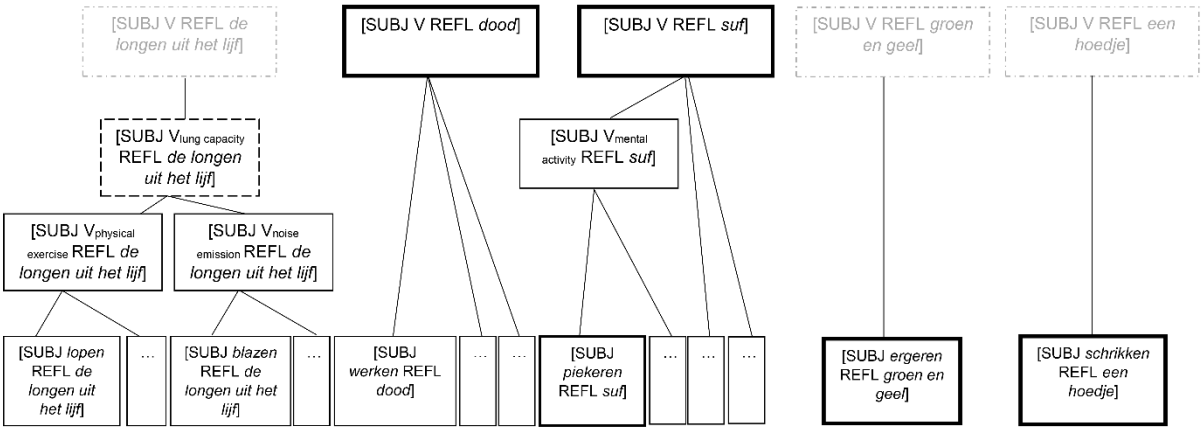


Figure 4.28. Towards building a constructional network, step 4

Figure 4.28 should give a relatively accurate representation of the lower hierarchic levels of the intensifier-centred network for these five intensifiers, insofar as we can infer from our data. Note that no semantically specified subschemas are posited in the verb-centred network, based on the data in SoNaR-NL. It was already pointed out in the previous section that verbs are generally less likely to impose collocational constraints on the intensifiers they pair up with. Of course, the fact that we did not find any obvious linguistic evidence for such subschemas in the SoNaR-NL data does not mean that semantically defined subschemas are by default impossible in a verb-centred network (as will be shown in §4.4.2 and Ch5, 5.4 infra). It may well be the case that language users perceive of certain similarities between the intensifiers that we, as linguists, did not notice or did not find to be particularly relevant.

We can now start generalising further upwards. We arguably still need to posit some intermediary abstractions before we arrive at the maximally schematic level of [SUBJ V REFL INT]. However, it is not exactly clear which kinds of broader generalisations the language user will make and to what degree of specificity these are formulated. It is possible that he/she generalises over the formal category of the intensifiers as in Figure 4.29. There are certain indications in our data that the syntactic category of the intensifier can indeed explain some of the differences in collocational behaviour and productivity: most of the NP+PP intensifiers are found to impose (more or less strict) semantic constraints, whereas the adjectival intensifiers are generally much more freely extensible to all kinds of verb types. It is also possible that the language user generalises over the semantics of the intensifier, as in Figure 4.30. While we have seen that semantics definitely play a role at the level of lower-level subschemas, it has been shown in 4.4 that the intensifiers with similar (original) lexical semantics do not necessarily show the same collocational behaviour. Once more, we argue that this is not a question of "either-or" but a question of "and-and". Language users are able to make both formal and semantic generalisations at the same time, even if both types of generalisations are not easily conflated into just one taxonomic representation of the network.

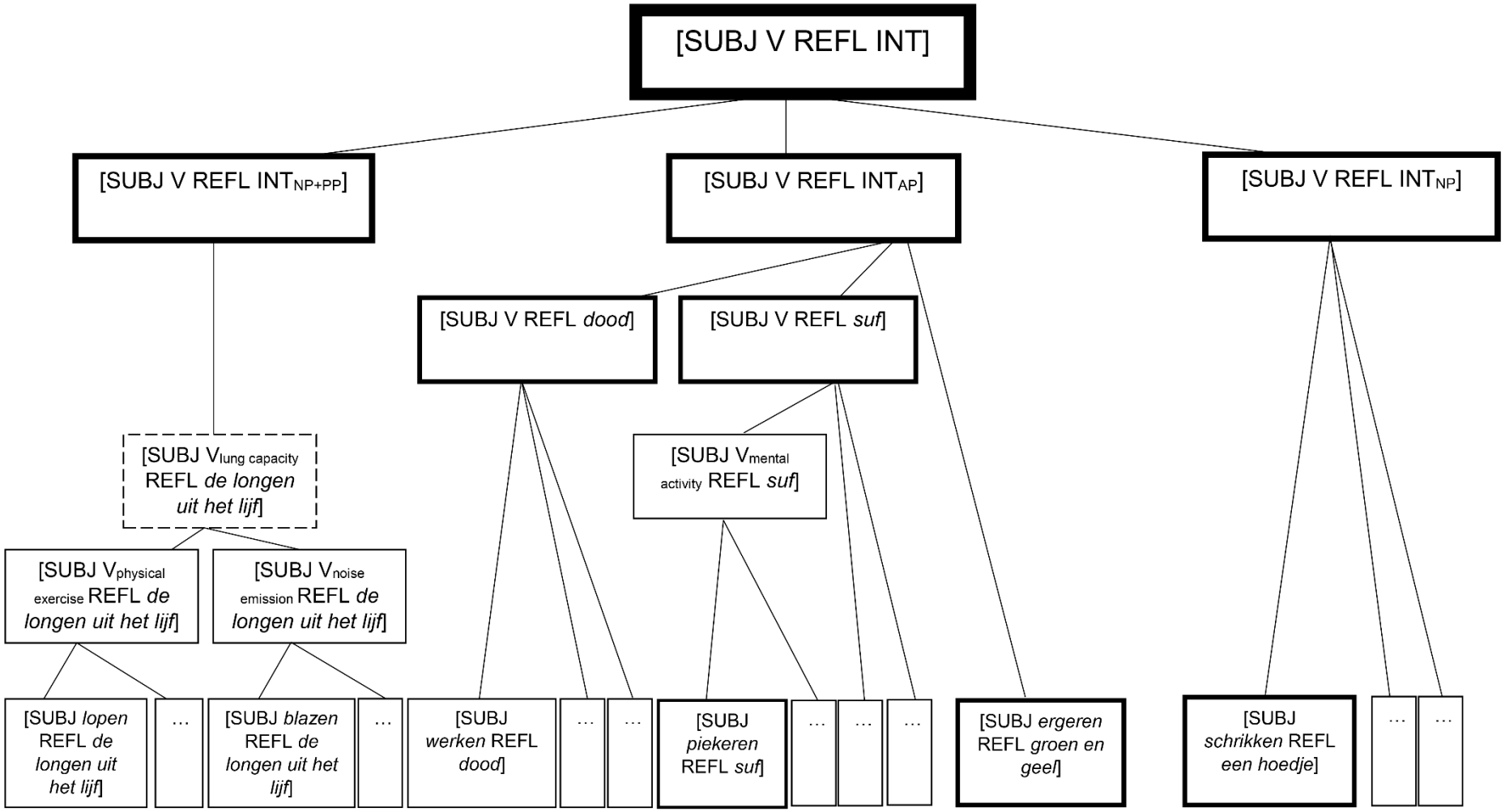


Figure 4.29. Towards building a constructional network, final step, INT syntactic category

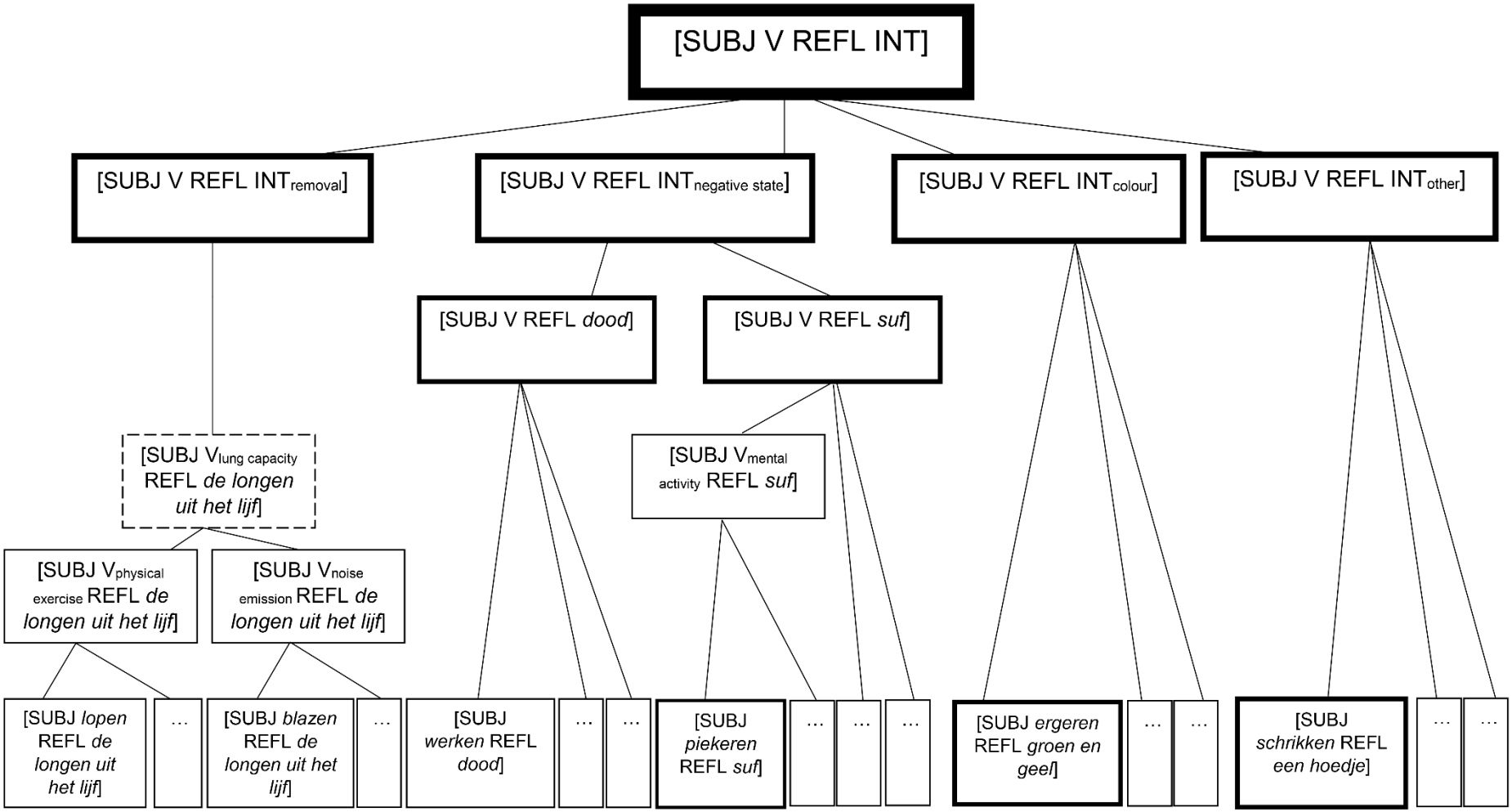


Figure 4.30. Towards building a constructional network, final step, INT semantics

For the verbs as well, we could opt for higher-order subschemas that are formally defined, based on the reflexivity or transitivity of the verb. Again, we emphasise that such a distinction is only useful if the data contain evidence that there are certain differences between these verbal categories that may give rise to different generalisations. While this was not found to be the case for the transitivity of the verb, we did observe that the reflexive versus non-reflexive verbs behaved slightly differently with respect to the distribution over the syntactic categories of the co-occurring intensifiers. For example, barely any of the NP+PP intensifiers or NP intensifiers were found to occur with inherently reflexive verbs (cf. §4.1.1.1 supra). Figure 4.31 distinguishes between the reflexive and non-reflexive verbs at a high level in the network, but it is not impossible that the distinction is only perceived as relevant at a much lower level in the network. It was also mentioned that it is not entirely clear from our data whether the behavioural differences that were observed are actually to be attributed to the formal property of reflexivity, or whether the semantics of the verb are of overriding importance. It appears that most of the frequent reflexive verbs like *zich ergeren* 'to be annoyed' or *zich vervelen* 'to be bored', as experience verbs, behave rather similarly to the non-reflexive experience verbs *lachen* 'to laugh' or *schrikken* 'to be startled'. In other words, Figure 4.32 is also a plausible representation of the verb-centred network. Note that this representation only captures very broad semantic categories, as we did not posit any fine-grained semantic classes in §4.1.1.1. We could of course further split up the experience verbs into emotional and cognitive experience verbs or make finer semantic distinctions within the group of physical activity verbs, if those differences were to be perceived of as relevant by the language user.

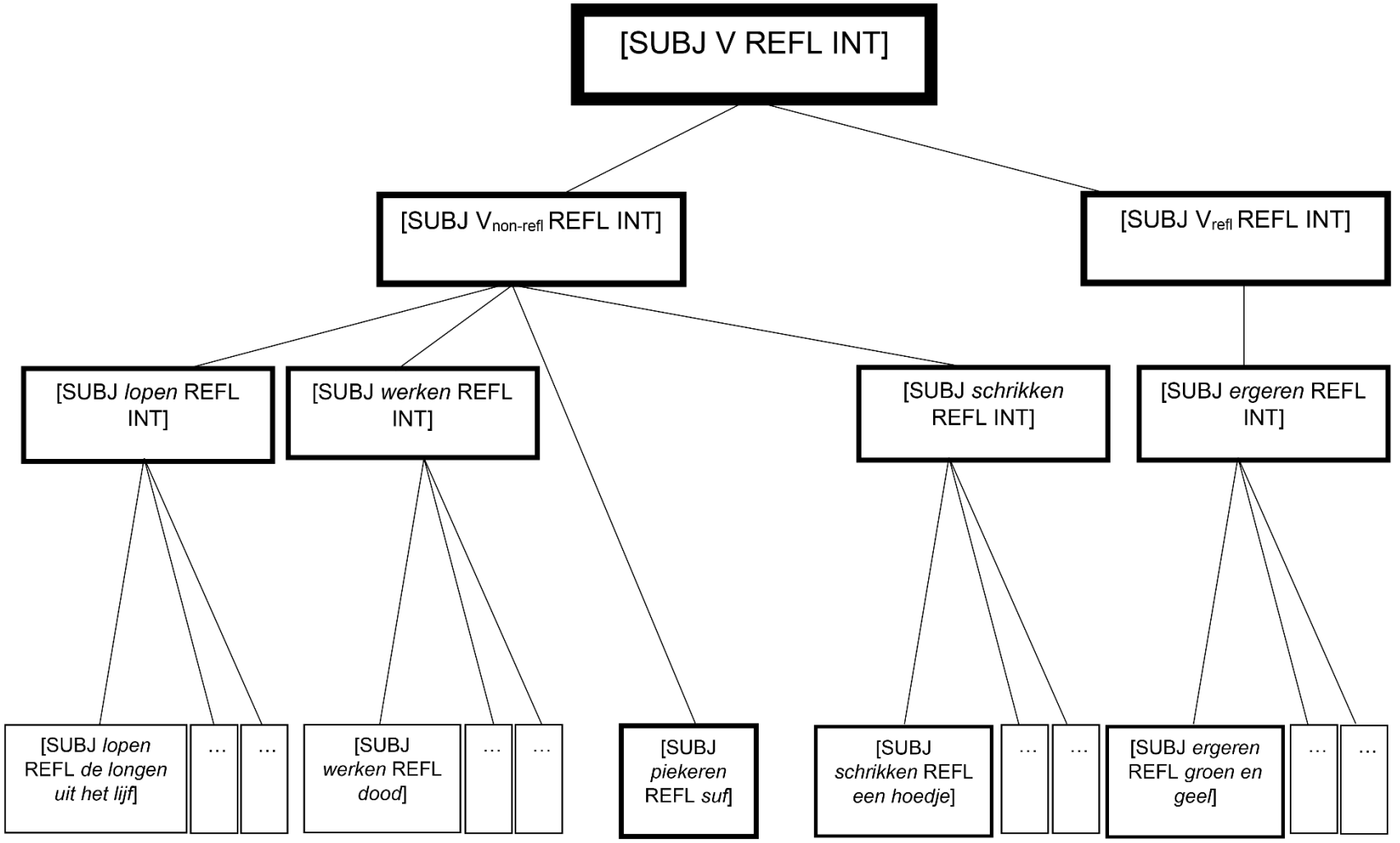


Figure 4.31. Towards building a constructional network, final step, V reflexivity

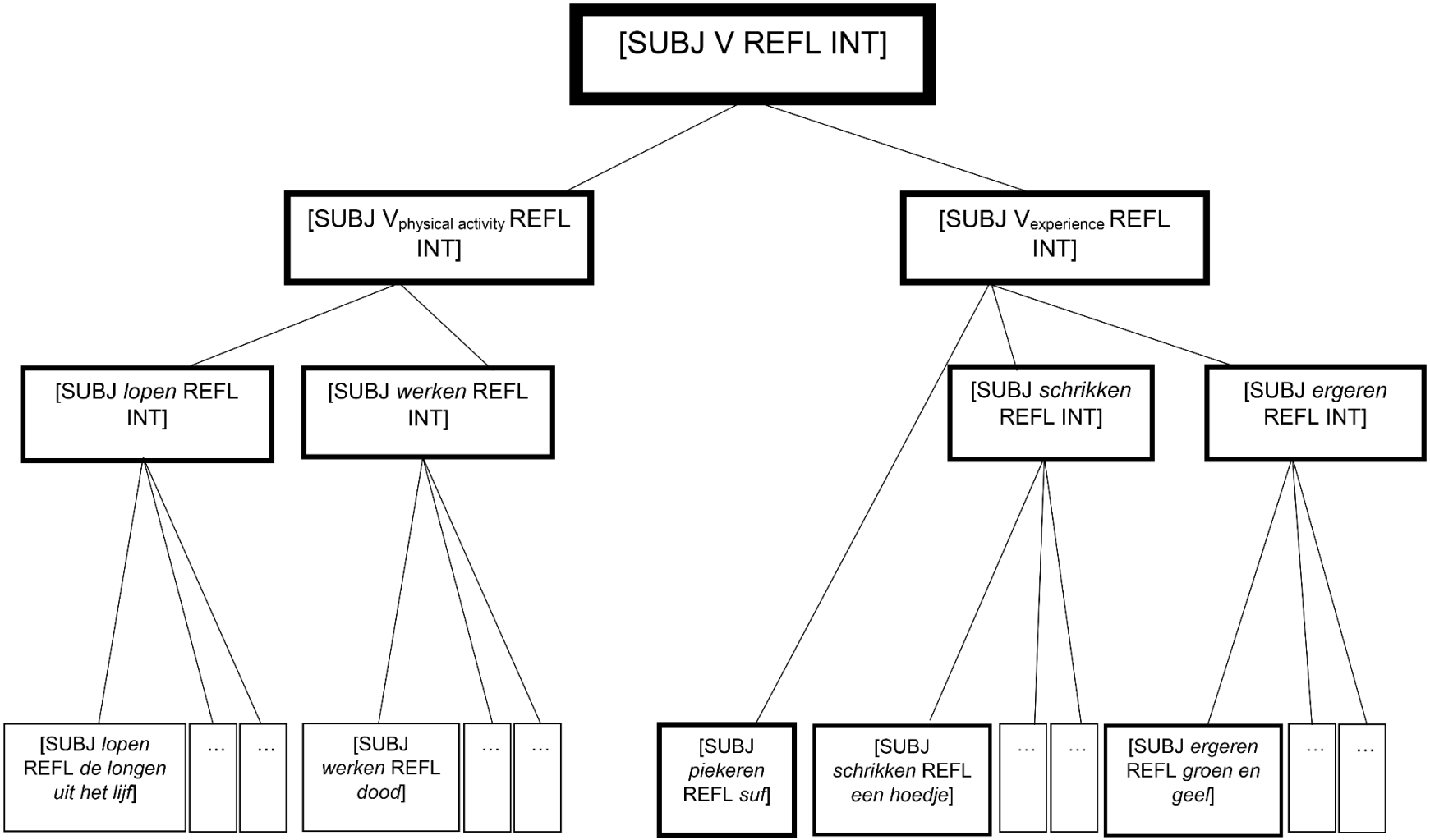


Figure 4.32. Towards building a constructional network, final step, V semantics

With Figure 4.29 to Figure 4.32, we have (for now) reached the final steps of the network building process. Of course these are just four possible representations of the constructional network of the intensifying fake reflexive resultative construction, based on the SoNaR-NL corpus data. We hope to have illustrated how both quantitative and semantic aspects of productivity factor into the representation of subschemas and micro-constructions at different levels in the network hierarchy. We deleted the "non-existent" subschemas in order to clearly visualise that the so-called isolated micro-constructions that did not trigger any intermediate generalisations are still motivated within the network because they are directly licensed by one of the higher levels in the hierarchy. In fact, all micro-constructions, regardless of their degree of entrenchment, are still instantiations of the more abstract schema [SUBJ V REFL INT] and are therefore always motivated by at least some other node in the network.

In general, though, most of the micro-constructions consist of either a verb or intensifier (or both) that has some combinatorial flexibility, and are thus subsumed by a lower-level subschema in at least one of the multiple possible network representations. Of the 214 out of 318 different verb-intensifier combinations that are one-offs (i.e. barely or not entrenched at all), there are only a couple of on-the-fly combinations of verbs and intensifiers that are not licensed by a lower-level verb-specific or intensifier-specific subschema. Given that both the verb and intensifier are infrequent, there is no sufficient basis to assume that the language user has formed a generalisation in the form of a partially lexically specified subschema. Examples like (196) and (197) below are therefore directly licensed by – and evidence for the existence of – a highly abstract schema in the higher regions of the network.

(196) De tweede vreest voor zijn privileges, maakt af en toe wat excuses, **manipuleert zich paars**. (SoNaR-NL)
[...] manipulates himself purple
'The second one is scared of losing his priviliges, comes up with excuses every now and then and manipulates the hell out of people.'

(197) Een jaar geleden **protesteerde** het volk **zich krankjorum**. (SoNaR-NL)
a year ago protested the people itself bonkers
'A year ago, the people held one protest after another.'

It was mentioned earlier in this section that the abstract schema [SUBJ V REFL INT] can be considered productive if we just look at the number of individual intensifier and verb types and the many hapaxes that can be filled into the slots. However, the fact that truly creative combinations, such as the ones in (196) and (197), are rather rare actually suggests that language users often do not just randomly put together verbs and intensifiers. It appears, then, that the productivity of the construction is not primarily determined by extensions at the highest level of schematicity, but that the real productivity of the intensifying fake reflexive resultative construction is situated at lower

levels in the network, such as the verb-specific and intensifier-specific subschemas. In Chapter 5, we will further illustrate the dynamics between the different levels in the hierarchy and demonstrate how the present-day network structure has been shaped and reshaped over the past two centuries. Chapter 6, §6.2.3 will return to the significance of measuring productivity and extension strategies at different levels in the hierarchy.

Although our focus has primarily been on taxonomic inheritance links, we briefly want to address how we could integrate horizontal links in the current representations of the network and what the theoretical status of these links are. In Chapter 2, it was mentioned that horizontal links have been invoked in the existing literature to capture (partial) similarities between patterns at the same level of abstraction that are not taxonomically licensed by or reducible to the same abstract schema. For example, Verhagen (2003b) mentions that the Dutch *weg*-construction as in *Hij baant zich een weg door de massa* 'He makes his way through the crowd' is tightly connected with other constructions like the causative construction with *weg* (e.g. *de weg banen/openen voor X* 'to pave the way for X') as well as the reflexive movement construction (e.g. *zich bewegen/worstelen/slepen door X* 'to move/struggle/drag oneself through X'). Evidence that language users capture such similarities across different constructions is found in priming experiments and L1-acquistion (see Diessel 2015 for some discussion), the phenomenon of constructional contamination (Pijpops & Van de Velde 2016, cf. 2.1.3), as well as in language change, with new constructions arising as an amalgam of two or more unrelated constructions (Hilpert 2014). Our first suggestion with respect to horizontal links in the intensifying fake reflexive resultative construction is somewhat inspired by these existing accounts in the sense that these horizontal links allow us to bring attention to similarities between micro-constructions (i.e. patterns at the same level of abstraction) that are not licensed by the same subschema and therefore not picked up in the taxonomic representation of the network. Given the numerous potential verb-intensifier combinations, there are evidently several intensifiers that share the same verbs in their collocational range. While this is captured by verb-specific subschemas in the verb-centred network, this similarity is not relevant in the hierarchic structure of the intensifier-centred network because we immediately abstract away from the specific verb. Instead, we can link these micro-constructions through horizontal links at the same level of abstraction, as in Figure 4.33.[53] Within a multirepresentational approach to network structure, we could say that these horizontal links in the intensifier-centred network correspond to the vertical links in the verb-centred network. This is a particular application of horizontal links that, to our knowledge, has not explicitly been treated as such in the literature.

[53] If we were to do this for all 318 micro-constructions, the visual representation would get immensely complicated, but virtual interactive models could account for both vertical and horizontal links in one and the same network representation.

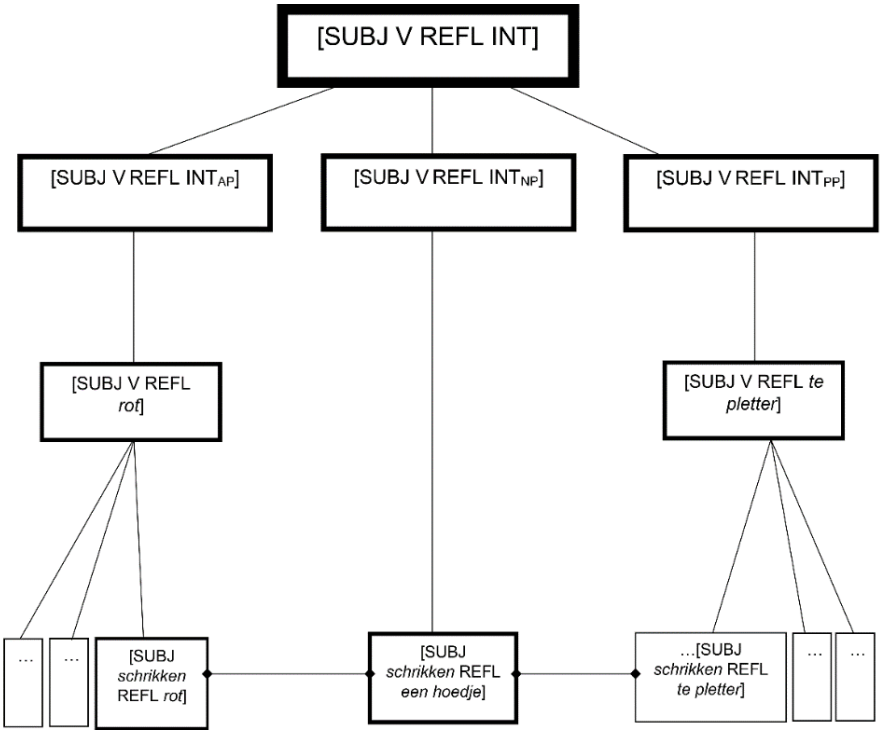


Figure 4.33. Horizontal links between shared verbs in the intensifier-centred network

Similarly, we can capture shared intensifiers at the micro-construction level in the verb-centred network as in Figure 4.34. Again, these horizontal links are reinterpretations of vertical links in a different representation of the network: in the intensifier-centred network, the micro-constructions would all be taxonomically motivated by the same subschema in which the intensifier *suf* 'drowsy' is lexically specified.

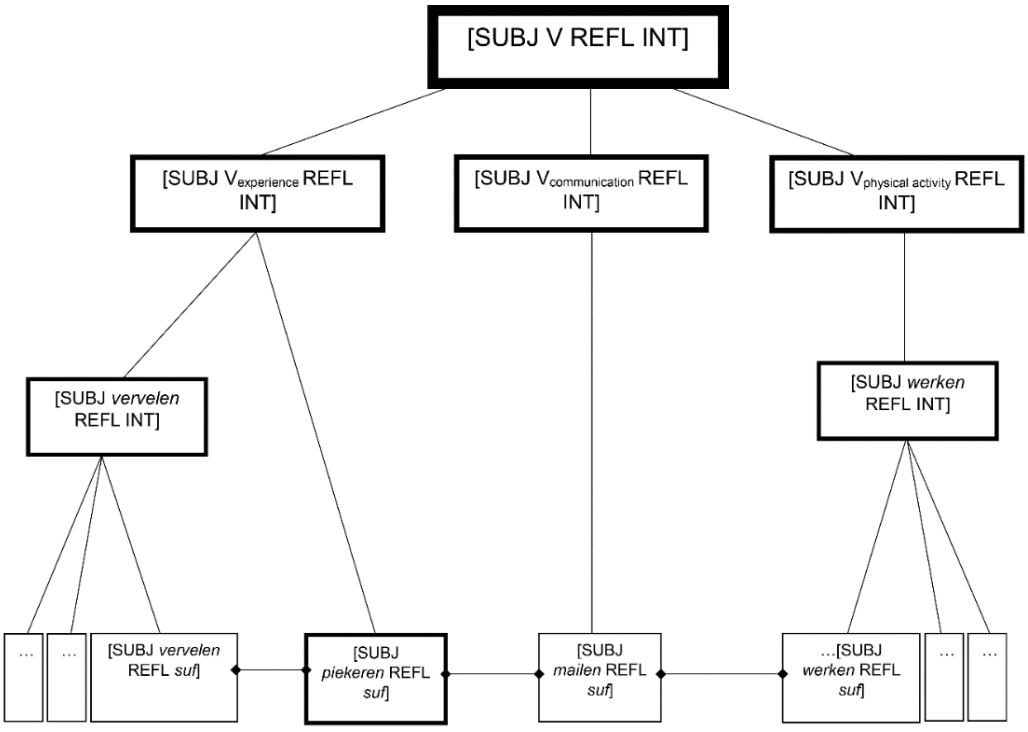


Figure 4.34. Horizontal links between shared intensifiers in the verb-centred network

These horizontal links also highlight that the subschemas at the same level of abstraction are in a kind of paradigmatic relationship, in the sense that they are interchangeable to a certain extent. The micro-constructions in the intensifier-centred network are in fact perfectly interchangeable without any important change in meaning: *zich rot schrikken* 'to startle oneself rotten', *zich een hoedje schrikken* 'to startle oneself a little hat' and *zich te pletter schrikken* 'to startle oneself to smithereens' all denote that someone is very startled. The verb-centred micro-constructions are also in a paradigmatic relationship, but the fact that they feature different verbs evidently changes the meaning. In the examples of *zich suf piekeren* 'to worry oneself drowsy', *zich suf mailen* 'to mail oneself drowsy' and *zich suf werken* 'to work oneself drowsy', *suf* 'drowsy' denotes that the verbal activities of worrying, mailing and working are performed with a certain intensity. The fact that our horizontal links may capture paradigmatic relationships makes them somewhat similar to those discussed by Van de Velde (2014). He suggests that constructions which form a kind of paradigm, i.e. "a set of alternating forms with related meaning differences", are related on a horizontal level (2014: 149). However, his paradigmatic relationships are different from the ones discussed here because they are strictly syntactic in nature. The examples he provides illustrate how alternating syntactic forms (i.e. forms which are in a syntactic paradigm) may contribute different meanings/functions. For example, the position of the verb in the main clause is linked to different kinds of clause types: V2 is found in main declarative clauses, V1 in polarity questions or imperatives and Vfinal in subordinate clauses. He also illustrates that the meaning differences expressed by horizontal links may survive if the network comes under pressure during language change.

A somewhat different (but also paradigmatic) kind of horizontal links can also provide a way of accounting for slight variations in the lexical content of a specific intensifier. For example, we already saw that in the collocation *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers', it is possible to replace *de sloffen* 'the slippers' by other types of footwear. While all of these can be considered as micro-constructions of the subschema [SUBJ *lopen* REFL INT] in the overall network structure, at the same hierarchic level as all other verb-intensifier combinations with *lopen* 'to run', they are still intuitively rather different from, e.g., *zich rot lopen* 'to run oneself rotten' or *zich suf lopen* 'to run oneself drowsy'. While this difference is not evident in the taxonomic structure of the network, we can highlight the tight connection to the model intensifier *het vuur uit de sloffen* 'the fire out of the slippers' through horizontal links, see Figure 4.35.

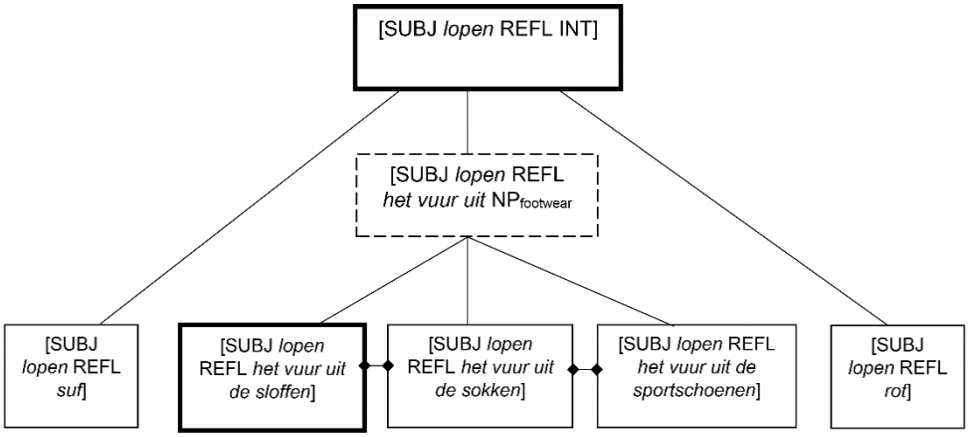


Figure 4.35. Horizontal analogical extensions

As these kinds of variations are primarily attested for highly token frequent collocations, they can be considered as prime examples of analogical extensions on the basis of a highly frequent model. In Chapter 5, it will be argued that such micro-constructions have originally entered the network as very local variations on a highly frequent model, rather than as new instantiations of a more abstract subschema – even though they, of course, came to be taxonomically licensed by that subschema as soon as they joined the network. If the language user is confronted with several of these analogical extensions, it is not implausible that these horizontal connections give rise to a superordinate subschema along the lines of [SUBJ *lopen* REFL *het vuur uit* $NP_{footwear}$]. In the next paragraph, we will briefly discuss a similar example for NP+*uit het lijf* intensifiers in Belgian Dutch.

4.4.2 Synchronic variation

The previous sections have highlighted a number of national differences in the way in which the individual slots of the construction are filled in Belgian and Netherlandic Dutch, but also in the way in which the verbs and intensifiers covaried or interacted with one another. Most of these differences were subtle variations in frequency rankings, intensifier preferences or lexical-collocational co-occurrences, but there were also more pronounced differences in the form of nationally exclusive – or at least highly preferred – intensifiers or verb-intensifier combinations. This section seeks to find out to what extent these differences between Belgian and Netherlandic Dutch have an impact on the taxonomic representations (and, perhaps, the cognitive organisation) of the constructional network of the intensifying fake reflexive resultative construction in both national varieties.

First of all, not all attested differences between the national varieties of Dutch are directly relevant for the hierarchic structure of the constructional network. There are certain properties of the construction that are part of the *cognitive* representation of the constructional network (in as far as the network should contain all knowledge that the

language user has about the intensifying fake reflexive resultative construction, cf. supra), but that do not immediately affect the *taxonomic* structure of the network as represented here. As our representation of the network is based on the current extent of use of the construction, the hapax-type and hapax-token ratios which are said to measure the extensibility of a pattern have little to no influence on the position of that pattern within the overall taxonomy. For example, the distance between the horizontal positions of *kapot* 'broken' (i.e. its potential productivity) in the global productivity graph in Belgian vs. Netherlandic Dutch (cf. Figure 4.17) does not really affect the place of the [SUBJ V REFL *kapot*] subschema in the taxonomy. Given that it occurs with a wide range of semantically diverse types in both national varieties, it is a moderately entrenched subschema in both network representations. Even the differences between the national equivalents of *te pletter* 'to smithereens' and *suf* 'drowsy' largely disappear in the hierarchic structure of the network. The intensifier *te pletter* is found with an array of verb types that do not show any obvious coherence in both national varieties, so the subschema [SUBJ V REFL *te pletter*] is positioned at the same level in both national networks. For *suf*, both national data sets contain evidence for a semantically unconstrained subschema [SUBJ V REFL *suf*], but also for a lower-level subschema specifying the verbs of mental activity and a strongly entrenched collocation with *piekeren* 'to worry' at the micro-construction level. However, there is of course a large difference in type frequency between both national varieties, with *te pletter* 'to smithereens' being much more type frequent in Belgian Dutch and *suf* 'drowsy' being extremely prolific in Netherlandic Dutch. As type frequency determines the degree of entrenchment of the subschema (just like token frequency or collostruction strength determine the entrenchment of the micro-construction), these differences should be represented in the network somehow. In other words, there are indications of differences with respect to the degree of entrenchment, i.e. the strength of the cognitive representation, of specific nodes in the constructional network, even though their hierarchic place in the overall taxonomy is not necessarily different. In some cases, the difference in entrenchment can be quite large. For example, [SUBJ *ergeren* REFL *groen en geel*] is strongly entrenched at the micro-construction level in Netherlandic Dutch, but much less so in Belgian Dutch, where it occurs only twice.

We further illustrate these differences in entrenchment by means of two interesting intensifiers which were not yet discussed in the previous paragraph, viz. *blauw* 'blue' and *uit de naad* 'out of the seam', both of which were found to be interesting in other respects as well (cf. §4.3.1.2 and §4.3.2.2). Starting with *blauw*, its status in the network is somewhat unclear in both national varieties. Based on the data, the subschema [SUBJ V REFL *blauw*] – if it exists at all – is probably not very strongly represented. Given its low scores for semantic coherence and type frequency and the high strength of the collocations with *zich ergeren* 'to be annoyed' and *betalen* 'to pay' (two verbs that show little semantic affinity), we might not have expected *blauw* 'blue' to be productive at all. Still, we need to find a way to account for the other verb types that are found in its collocational range in

both national varieties. Of course, we could assume some kind of interaction with the verb-centred network, in which those instances of *blauw* 'blue' are motivated by the respective verb-specific subschemas, just like we did in §4.4.1 for *zich een hoedje lachen/schrikken* 'to laugh/startle oneself a little hat' and *zich rot/suf piekeren* 'to worry oneself rotten/drowsy'. However, as most of these other verbs occurring with *blauw* 'blue' are themselves rather infrequent, the micro-constructions may not be motivated by a lower-level subschema in the verb-centred network either. It therefore appears that they are examples of non-entrenched micro-constructions that are immediately licensed by a higher-level abstract schema (i.e. instances that were all in all found to be rather rare, cf. supra). Note that even if the micro-constructions with *blauw* 'blue' are not subsumed by a low-level subschema, we can highlight their shared intensifier by linking them on a horizontal level (cf. §4.4.1 supra). While Figure 4.36 gives the same network structure for both the Belgian and the Netherlandic Dutch variants of *blauw* 'blue', there is some difference in the degree of entrenchment of the micro-constructions. Given that the collocations *zich blauw ergeren* and *zich blauw betalen* are stronger in Belgian Dutch (cf. §4.2), the micro-constructions are even more entrenched in the Belgian network than in the Netherlandic network, as indicated by the grey shade in the left panel of Figure 4.36.

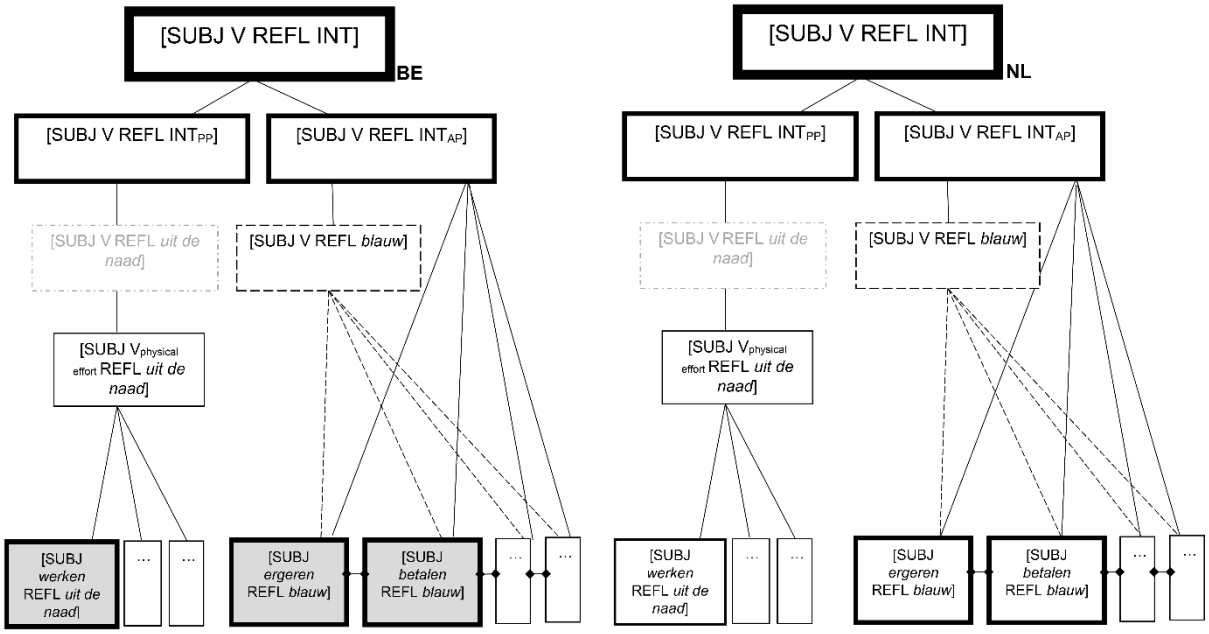


Figure 4.36. National differences in the constructional network, entrenchment

Uit de naad 'out of the seam' was one of the intensifiers that displayed obvious semantic coherence among its types, insofar as all the verbs occurring in the verb slot involve some physical effort. This means that the subschema [SUBJ V REFL *uit de naad*] is most likely too general (cf. the grey dashed lines) and that a lower-level generalisation in the form of [SUBJ $V_{\text{physical effort}}$ REFL *uit de naad*] is more faithful to the data. Although *uit de naad* 'out of the seam' also has one attestation with a noise emission verb in both national varieties, this category is likely not sufficiently prominent to give rise to a separate intermediate subschema (unlike *de longen uit het lijf* 'the lungs out of the body', cf. supra, or *de ziel uit het*

lijf ‘the soul out of the body’, cf. infra). Again, the taxonomic structure looks identical in both national varieties. Within the physical effort subschema, we find one entrenched micro-construction [SUBJ *werken* REFL *uit de naad*] that accounts for about 70% of all the tokens with *uit de naad* ‘out of the seam’ in both national varieties. It is not unlikely that *zich uit de naad werken* ‘to work oneself out of the seam’ was once a highly frequent fixed expression which came to serve as an analogical model that, through local extensions, gave rise to the partially productive subschema [SUBJ $V_{\text{physical effort}}$ REFL *uit de naad*]. The diachronic development of *uit de naad* ‘out of the seam’ will be investigated in Chapter 5. We highlight the special synchronic status of this collocation by marking it as entrenched. Looking back at the results of the covarying collexeme analysis in §4.2, we see that the collocation is much higher ranked in Belgian Dutch than in Netherlandic Dutch so we indicated this stronger level of entrenchment by adding grey shades in the left panel of Figure 4.36.

More interestingly with a view to network organisation, we have also found a small number of overlapping intensifiers which display a different degree of combinatorial flexibility in both national varieties. For example, we found that the 13 tokens of *de pleuris* ‘the pleurisy’ in Belgian Dutch contained 7 different verb types (4 of which are hapaxes), whereas in Netherlandic Dutch, *de pleuris* occurs only 3 times, twice with *schrikken* ‘to be startled’. The exact opposite is found for *wezenloos* ‘vacant’, which is found only once in Belgian Dutch, but the 22 occurrences of which in Netherlandic Dutch feature 9 semantically diverse verb types (6 of which are hapaxes). These differences do have repercussions for the overall structure of the network in Figure 4.37: in Belgian Dutch, there appears to be a somewhat productive subschema [SUBJ V REFL *de pleuris*] that is absent in Netherlandic Dutch, whereas Netherlandic Dutch has a productive subschema [SUBJ V REFL *wezenloos*], which is not represented in the Belgian network.

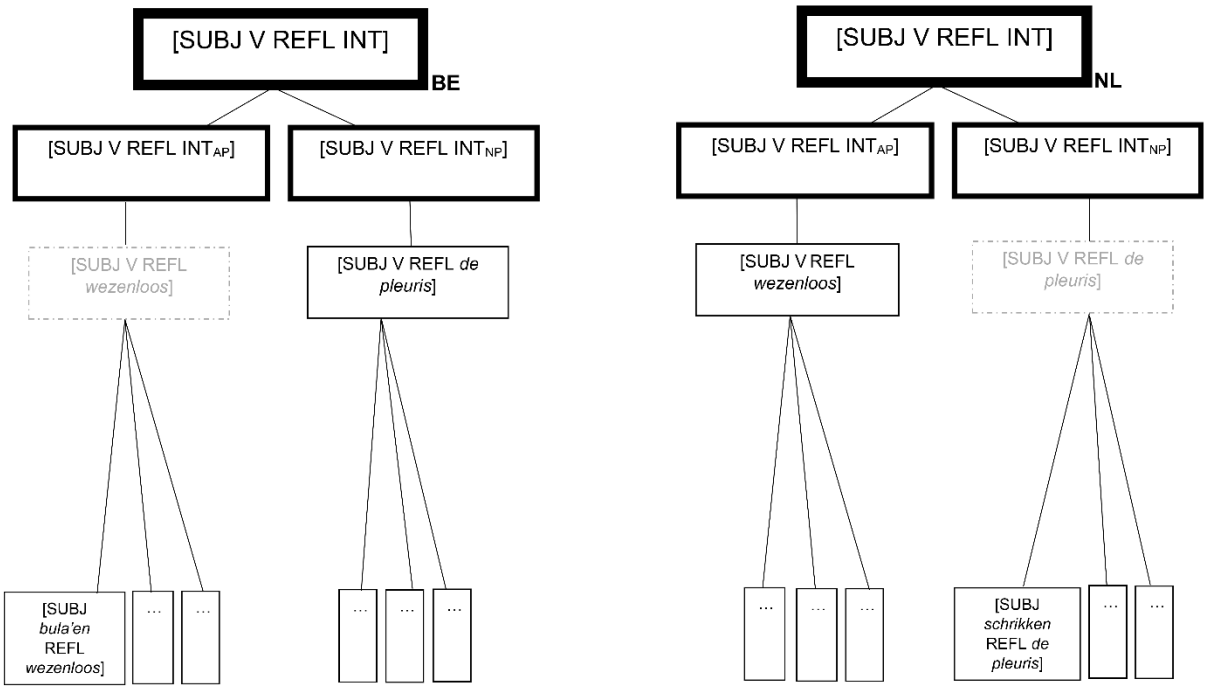


Figure 4.37. National differences in the constructional network, schematicity

Evidently, we also need to account for the non-overlapping intensifiers in the overall network structure. In previous sections, we already argued that not all of these non-overlapping intensifiers display the same degree of productivity, implying that they should be represented at different levels in the network. There are only a couple of nationally exclusive intensifiers that could be represented at the subschema level. From the top 15 intensifiers, this only applies to *de ziel uit het lijf* 'the soul out of the body' in Belgian Dutch and *een slag in de rondte* 'a punch around' in Netherlandic Dutch. We found that *de ziel uit het lijf* 'the soul out of the body' imposes exactly the same semantic constraints on its verb slot as *de longen uit het lijf* 'the lungs out of the body', so the taxonomic structure also looks virtually identical in Figure 4.38. For *de longen uit het lijf* 'the lungs out of body', we suggested that the language user may even generalise over the physical exercise verbs and the noise emission verbs in a higher-level schema with the restriction $V_{\text{lung capacity}}$. However, this constraint was admittedly posited as rather ad-hoc, given the lexical element of *longen* 'lungs': as this element is not present in *de ziel uit het lijf* 'the soul of the body', it may be rather odd to define a subschema with $V_{\text{lung capacity}}$ here. While the fact that the physical exercise verbs and noise emission verbs can occur with the same type of intensifiers does suggest a certain semantic relatedness between these categories, the shared element may not actually be directly related to lung capacity after all. A better definition could be something more neutral as "physical effort" (cf. *uit de naad* 'out of the seam', supra). This shows that the linguist must be careful not to over-interpret the data and remain aware that corpus data may only provide a partial insight into the kinds of generalisations that are relevant to language users.

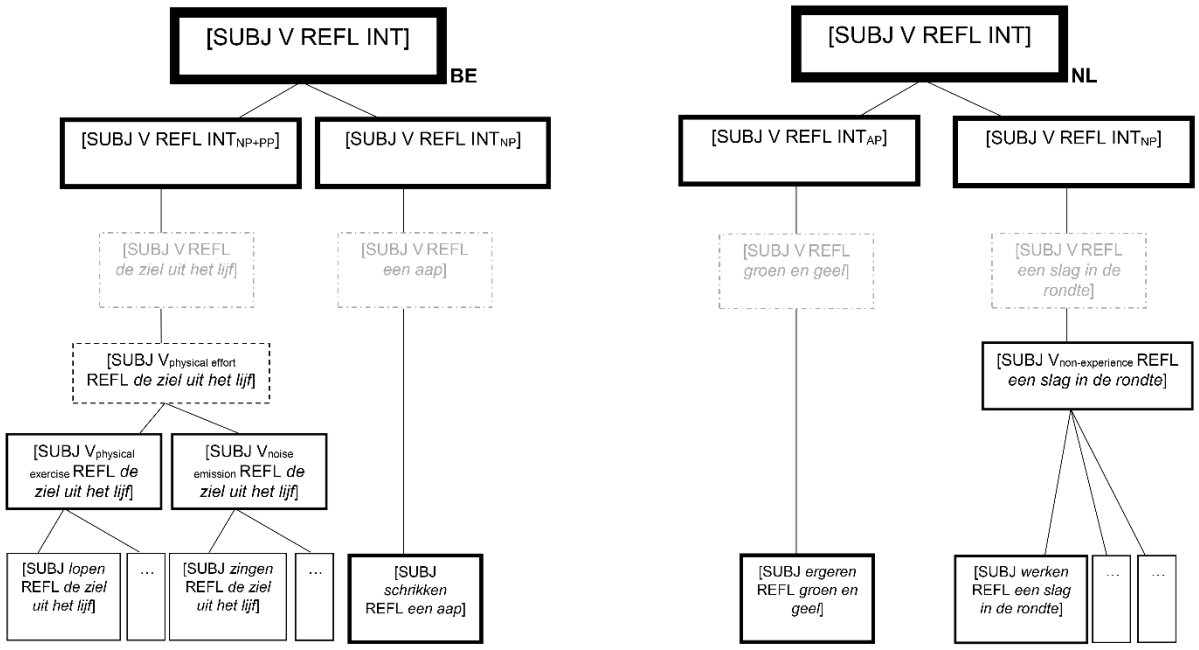


Figure 4.38. National differences in the constructional network, non-overlapping intensifiers

Of course, it is possible that the language user also perceives of shared lexical elements between intensifiers. If so, the representation of the network in Belgian Dutch may require an extra subschema somewhere that captures the fact that *de longen uit het lijf* 'the lungs out of the body' and *de ziel uit het lijf* 'the soul out of the body' (as well as *de naad uit het lijf* 'the seam out of the body', which also has more than 10 attestations in SoNaR-BE) share the PP *uit het lijf* 'out of the body', viz. [SUBJ V$_{\text{physical exercise/noise emission}}$ REFL *de* N *uit het lijf*]. The hapax intensifier *de balg uit het lijf* 'the gut out of the body', see (198), would also be subsumed by that subschema, while at the same time being horizontally linked to one or several of the more frequent models (cf. the examples of *het vuur uit de* N$_{\text{footwear}}$ in §4.4.1).

(198) Dat beest zit vast klem tussen de keien van de rotspartij rond het vijvertje en **schreeuwt zich de balg uit het lijf** om verlost te worden. (SoNaR-BE)
[...] and screams itself the gut out of the body [...]
'The animal is probably stuck between the rocks around the pond and is screaming its gut out to be released.'

It is possible that some of the currently more frequent NP+*uit het lijf* 'out of the body' intensifiers have originally arisen as (infrequent) analogical extensions, much like *de balg uit het lijf*, but over time have come to attract more and more verb types and developed into separate subschemas. As we do not have any diachronic Belgian data, we cannot say which of the intensifiers was the "original" model or at what time the others were created, but some similar cases for Netherlandic Dutch will be discussed in Chapter 5.

The Netherlandic intensifier *een slag in de rondte* 'a punch around' is slightly more difficult to position. Its verbal collocates do not really belong to one semantically coherent class, but we did point out that one semantic class was surprisingly absent, viz.

the experience verbs. We can integrate this in the network by adding a *negative* constraint in the subschema, as in Figure 4.38. The micro-construction [SUBJ *werken* REFL *een slag in de rondte*] is indicated as a moderately entrenched micro-construction because, even though it was not part of the top twenty in the covarying collexeme analysis in §4.2.1, it is still a strong collocation at rank 22, accounting for 14 out of 37 tokens.

The final nationally exclusive intensifier that we will discuss is the Belgian intensifier *een aap* 'a monkey'. Given that it forms a near-exclusive collocation with the verb *schrikken* 'to be startled' in Belgian Dutch, language users probably do not form a generalisation in the form of [SUBJ V REFL *een aap*]. Rather, [SUBJ *schrikken* REFL *een aap*] is an entrenched micro-construction in the Belgian network (rank 9 in the covarying collexeme analysis in §4.2.2) that is immediately licensed by a higher-order abstract subschema in the intensifier-centred network, just like [SUBJ *ergeren* REFL *groen en geel*] in the Netherlandic network (cf. supra).

So far, we have only considered the intensifier-centred network representations. From the perspective of the verb-specific subschemas, most of the differences relate to the degree of entrenchment at the micro-construction level. For example, the subschema [SUBJ *ergeren* REFL INT] has (at least) two highly entrenched micro-constructions [SUBJ *ergeren* REFL *wild*] and [SUBJ *ergeren* REFL *groen en geel*] in Netherlandic Dutch that are not present or not well-entrenched in the Belgian network; conversely, in the Belgian network we find entrenched micro-constructions of [SUBJ *schrikken* REFL INT] like [SUBJ *schrikken* REFL *een aap*] and [SUBJ *schrikken* REFL *een bult*] that are not represented or not well-entrenched in the Netherlandic network. As was mentioned earlier, differences in the proportion of hapaxes or potential productivity do not directly influence the *taxonomic* representation of the network, but they may hint at a potential difference in the level at which certain schemas should be positioned in the network. If we look back at Table 4.12 and Table 4.16, we see that *zich schamen* 'to be embarrassed' was very token frequent in Netherlandic Dutch (80 tokens), but that it only occurred with 4 verb types and zero hapaxes. The absence of infrequent or one-off instances could be interpreted as the absence of a superordinate subschema [SUBJ *schamen* REFL INT]. Instead, *zich schamen* is part of four entrenched micro-constructions – the exact level of entrenchment varying according to the strength of the collocation, hence the slight difference in thickness in Figure 4.39. In Belgian Dutch, in contrast, the 19 tokens of *zich schamen* are distributed over 5 verb types. With the exception of *dood* 'dead', which is found in 12 of the 19 tokens, we find four infrequent intensifiers from the same semantic domain, viz. *te pletter* 'to smithereens', *rot* 'rotten', *steendood* 'stone-dead' and *kapot* 'broken' (the combination with *de ogen uit het hoofd* 'the eyes out of the head' is not found in Belgian Dutch). That is to say, the Belgian data do contain some evidence that the language users form a generalisation in the form of [SUBJ *schamen* REFL $\text{INT}_{\text{negative state}}$].

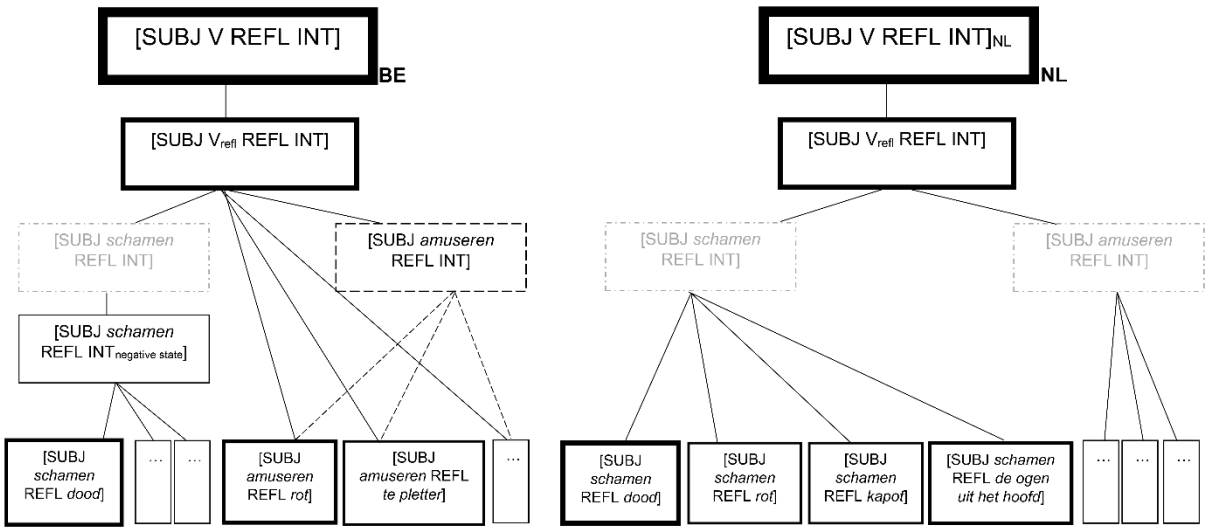


Figure 4.39. National differences in the constructional network, verbs

There was also a large discrepancy in token frequency for the verb *zich amuseren* 'to enjoy oneself'. With only 5 tokens and 3 types in Netherlandic Dutch, a verb-specific subschema is rather unlikely. In Belgian Dutch, there were 75 tokens but only 5 types. Especially the micro-constructions [SUBJ *amuseren* REFL *rot*] and [SUBJ *amuseren* REFL *te pletter*] are well-entrenched, according to the covarying collexeme analysis. We could interpret the remaining types as evidence of the existence of the subschema [SUBJ *amuseren* REFL INT], but they could also be licensed directly by a higher-order schema, parallel to what we suggested for the intensifier *blauw* 'blue'. As mentioned earlier, we do not know how many different types are needed in order for the language user to make a generalisation and the kind of generalisations may vary according to speaker, context and specific verb/intensifier.

Finally, we found that there is apparently some kind of association between the verb *zich ergeren* and colour intensifiers. We may need to posit an intermediary generalisation in the form of [SUBJ *ergeren* REFL INT$_{colour}$] in between [SUBJ *ergeren* REFL INT] and the concrete micro-constructions [SUBJ *ergeren* REFL *blauw/groen en geel/paars/...*]. As this association is especially obvious in the Belgian data, given that there are more different types of colour adjectives, the current corpus data suggest that the subschema should perhaps only be represented in the Belgian network. In Chapter 5, however, we will see that the preference of *zich ergeren* 'to be annoyed' for colour terms, although not particularly obvious in the SoNaR-NL data – starts to emerge in Netherlandic Dutch from the mid 20th Century and was still very prevalent in the 1990s. It is possible that *zich ergeren* 'to be annoyed' has since then retreated to the more conventional colour intensifiers in Netherlandic Dutch – but not (yet) in Belgian Dutch – , but it is not excluded that the absence of other colour intensifiers in the present-day Netherlandic data is due to some quirk in the corpus.

In sum, this paragraph has demonstrated that we should at least consider adding a national-variational dimension to constructional network organisation. Granted, there

are a lot of rather subtle variations between Belgian and Netherlandic Dutch that have little to no impact on the network organisation, but we also highlighted a number of more substantial differences with respect to the level at which certain subschemas or micro-constructions are situated in the network.

4.5 Interim conclusion

4.5.1 The intensifying fake reflexive resultative construction in present-day Netherlandic Dutch

In this chapter we analysed the synchronic use of the intensifying fake reflexive resultative construction in present-day Dutch by focusing on four main aspects of the construction. In the first section, we provided a general descriptive analysis of the individual slots of the construction. The intensifying data set contained 1,042 tokens, featuring 137 different verb types and 68 different intensifier types. The verb slot was found to allow for an extremely wide variety of different verb types from different semantic classes. Still, some verbs – especially the experience verbs and the verbs denoting physical activities – turn out to be more predisposed to being boosted in this specific construction than others. In that regard, we also discovered that the inherently reflexive verbs play an important role in the construction, in that just three inherently reflexive verbs, *zich ergeren* 'to be annoyed', *zich schamen* 'to be embarrassed' and *zich vervelen* 'to be bored', already account for a quarter of the entire data set. The intensifier slot showed remarkable variation as well, both in terms of its syntactic properties and its semantic properties. We already saw in Chapter 2 that the intensifiers can take different syntactic shapes, viz. adjectival phrases, nominal phrases and prepositional phrases, as well as combinations of a nominal and a prepositional phrase or (rarely) a nominal and an adjectival phrase. The adjectival category was found to be best represented overall, but the nominal category showed the highest degree of lexical variation with the highest proportion of different intensifier types. A wide array of intensifiers were recruited from a limited number of semantic source domains, most of which had some inherently negative connotation. It is this inherent negative aspect that may have contributed to their being used as an intensifier in the construction. At the same time, there are a number of intensifiers that did not fit into one of the delineated domains. The latter category does not only include infrequent (creative) creations but also some conventional intensifiers like *een hoedje* 'a little hat', the intensifying origins of which are much less obvious than for, e.g., *dood* 'dead'. Given that a small number of highly frequent intensifiers already account for a large part of the data set, it appears that conventionality

plays an important role in the use of the construction. At the same time, we find a large group of very infrequent or one-off intensifiers, attesting to the creative potential of the construction. While these hapaxes are often semantically related to one of the more frequent intensifiers, there are several hapax intensifiers that appear to be deliberately unconventional to create a special rhetorical effect.

In the second section, we studied the collocational patterns that the verbs and intensifiers enter into. A hierarchical cluster analysis was used to identify groups of verbs and intensifiers that show similar collocational behaviour. The suggested clusters did not entirely align with the semantic classes that were identified in section §4.2, in that most of the clusters were semantically heterogeneous or "incomplete", in the sense that at least one intensifier that should have been part of the cluster (based on its semantics) ended up in another cluster. A plausible explanation is that the original semantics of the intensifier have bleached to such an extent that they no longer exert much influence on the collocational behaviour. This may have led to collocational expansion and, accordingly, highly flexible combinatorial behaviour. Of course, it is also possible that semantically similar items have undergone different collocational *specialisation* (rather than expansion) and have developed idiosyncratic patterns. Given the 137 different verb types and 68 different intensifier types, there are over 9,000 potential verb-intensifier interactions, but the actual number of verb-intensifier combinations is "only" at 318. It was shown that the verb slot and the intensifier slot are generally not instantiated independently, but that the lexical filler of one slot covaries with the lexical filler of the other slot. There is a great deal of conventionality involved in this covariation: there are several strong verb-intensifier collocations that occur much more frequently than we would expect if the verbs and intensifiers were paired up at random and there are also many potential combinations that do not occur at all.

The discussion of collocational behaviour seamlessly brings us to the discussion of productivity in the third section of this chapter. This section has illustrated why productivity should be regarded as a multifaceted phenomenon. We can look at a pattern's type frequency to get an idea of its current extent of use or its "past achievement", but this fails to give us any information on the likelihood of the pattern being extended to new types. The extensibility or productivity of a pattern is determined by a number of elements. A central role was attributed to the hapaxes, a high proportion of which (among tokens or types) may indicate that the pattern is not associated with a number of (highly) frequent types. In that regard, we have suggested that the shape of the distribution can also be informative, but the measure of relative entropy was dismissed for being unpredictable, given that its implications for productivity have not been sufficiently verified. In addition, the semantic constraints that pertain to a specific slot in the pattern may limit the combinatorial flexibility and extensibility of a schema in important ways. This is why we have stressed the importance of integrating both the frequency-based measures, that take into account token frequencies, type frequencies

and hapax counts, as well as the constructional model that integrates frequency and semantic information into a truly multidimensional model of productivity. Broadly speaking, we find that the frequency-based measures by Baayen and colleagues and the productivity model by Barðdal lead to similar "productivity rankings", in that the subschemas that had relatively low scores for the frequency-based measures also appear to be on the lower end of the productivity cline, whereas the subschemas with high productivity scores are situated near the top end of the cline. However, the assumed inverse correlation between type frequency and semantic coherence was not always confirmed by the data, which shows that the model may need to be slightly refined (at least for this specific type of constructions).

Importantly, the analyses in the third section have also shown that it is crucial to differentiate between productivity at different levels of abstraction or schematicity. The taxonomy of the constructional network was further investigated in the fourth section. The aim of this section was to build a possible representation of the constructional network of the intensifying fake reflexive resultative construction, but we have called into question the reality of just one "final representation" of the constructional network. It is much more plausible that language users simultaneously have access to different, interacting configurations of the network, depending on the generalisations they make over multiple elements of the construction. While we may not be able to visually represent this complex cognitive network, we have illustrated the value of such a multiconfigurational or multirepresentational approach to network structure in explaining certain peculiarities in the use of the intensifying fake reflexive resultative construction. In general, the wide array of verbs and intensifiers, as well as the occasional disregard of collocational conventions was interpreted as evidence for the existence and productivity of the pattern at the most schematic level [SUBJ V REFL INT]. Still, it appears that the use of the construction is mainly determined by productivity at lower levels in the hierarchy. The network of the intensifying fake reflexive resultative construction consists of multiple pockets of productivity or productivity islands at different intermediate levels of schematicity. In some cases, a frequently occurring verb-intensifier combination, i.e. a lexically specified micro-construction, can be equally or more entrenched than the higher-level schema. In other cases, no superordinate low-level subschema can be assumed at all, which is why we find a number of scattered micro-constructions in the network. Of course, even these micro-constructions are at the very least motivated by a more abstract pattern at a much higher level in the hierarchy. Moreover, if we assume that different representations of the network may interact with one another, a micro-construction that appears to be isolated in one network representation may receive additional motivation by a low-level schema in another representation of the network. Finally, we have introduced horizontal links into the network in order to account for similarities between micro-constructions that are not subsumed by the same subschema in the taxonomic hierarchy. Summing up, we could

state that the intruiging mix of productivity and convention that is characteristic of this construction is related to the organisation of its constructional network, which is made up of both open, productive subschemas and lexically specified micro-constructions.

4.5.2 Constructional variation in national varieties of Dutch

In the second part of each subsection, we compared (i) the general usage, (ii) the collocational patterns, (iii) the productivity at different levels of abstraction and (iv) the taxonomic organisation of the constructional network of the intensifying fake reflexive resultative construction in both national varieties of Dutch. Overall, the analysis revealed numerous commonalities between the use of the construction in both national varieties, so this section will summarise the main differences.

First of all, the construction is more frequently attested in journalistic language in Netherlandic Dutch, with 142.5 instances per ten million words, than in Belgian Dutch, with 105.2 instances per ten million words. We then had a closer look at the individual slots to see whether the national varieties also show differences with respect to the concrete instantiation of the construction. Despite an overall lower variety of different verb types in Belgian Dutch and some frequency discrepancies of a couple of individual verbs, the verb slot did not reveal any notable differences between the national varieties of Dutch. With respect to the intensifier slot, it appears that Netherlandic Dutch also allows for a wider lexical variety of intensifiers (68/1,042) than Belgian Dutch (96/2,445), but the proportion of hapaxes is larger in Belgian Dutch (approximately 48%) than in Netherlandic Dutch (approximately 34%). Given that hapaxes in a corpus are assumed to reflect the creative uses of a construction, this would indicate that Belgian speakers (or journalists) are more creative in their use of this construction than Netherlandic speakers/journalists. Overall, there is a substantial overlap of 41 intensifiers occurring in both national varieties. The fact that 14 of the overlapping intensifiers are significantly more frequent in one of the two varieties indicates that there are different preferences with respect to the use of specific intensifiers in Belgian and Netherlandic Dutch. Most of the *non*-overlapping intensifiers were found to be rather infrequent, so we could not say whether those are examples of truly nationally exclusive intensifiers. Nonetheless, there were a number of intensifiers with more than 10 occurrences in one variety, while being altogether absent in the other. Exclusively used in Netherlandic Dutch were *een slag/slagen in de rondte* 'a punch/punches around', *wild* 'wild' and *ongans* 'unwell'; in Belgian Dutch, we find *de ziel uit het lijf* 'the soul out of the body', *een aap* 'a monkey', *steendood* 'stone-dead', *een beroerte* 'a stroke', *zot* 'crazy' and *de naad uit het lijf* 'the seam out of the body'. We need to add some nuance to the so-called "nationally exclusive intensifiers". It appears that in some cases, it is the entire verb-intensifier collocation that is nationally exclusive, suggesting that speakers of Belgian and Netherlandic Dutch have

formed different idiomatic expressions. In other words, rather than claiming that *een aap* is a Belgian-exclusive intensifier, we should say that *zich een aap schrikken* 'to startle oneself a monkey' is a Belgian expression. Similarly, *wild* is not a Netherlandic-exclusive intensifier, but *zich wild schrikken* 'to startle oneself wild' and *zich wild ergeren* 'to annoy oneself wild' are Netherlandic conventional collocations. We also found some evidence of national conventionalisation effects for overlapping intensifiers. For example, while *zich groen en geel ergeren* 'to annoy oneself green and yellow' is also attested in Belgian Dutch, it is a much more frequent collocation in Netherlandic Dutch, taking the top spot in the covarying collexeme analysis. In order to assess the impact of the attested national variation on the taxonomic representation(s) of the constructional network in Belgian and in Netherlandic Dutch, we first examined the differences in the degree of productivity of the subschemas that should (or should not) be included in the constructional network. The application of our multidimensional productivity model to the Belgian data revealed that many of the top overlapping intensifiers and verbs were about equally productive in both national varieties. Still, there were some differences with respect to the realised and/or potential productivity of individual verbs or intensifiers, as was obvious from the rather large distance between the national equivalents on the global productivity graph. The frequency discrepancies between the overlapping items primarily impacted the degree of entrenchment, although we did find some evidence of national variation in the overall taxonomic structure as well – especially if we add the non-overlapping intensifiers to the fold. To be fair, the overall structure of the network is largely the same in the two national varieties of Dutch, but to understand some of the intricacies of language use we need to focus our attention on the lower levels of the network, paying particular attention to lexically idiosyncratic uses of specific patterns and local motivations.

The main aim of this chapter was to provide a descriptive overview of the synchronic use and national variation of the intensifying fake reflexive resultative construction in present-day Dutch. In light of the increasing attention to lectal variation within the framework of Construction Grammar (cf. Chapter 2), we hope to have demonstrated that national variation may also have an impact on the productivity of a construction and the organisational structure of its constructional network. Part of the explanation for these national differences may be found in the linguistic history of the national varieties of Dutch (cf. supra). Other studies into lexical and syntactic variation in Belgian and Netherlandic Dutch, some of which were mentioned in §4.2.1 above, have found that register differences are generally more pronounced in Belgian Dutch than in Netherlandic Dutch, which shows a much more uniform distribution across different registers. This conservation of register differences in Belgian Dutch has been related to its delayed standardisation (see, e.g., Geeraerts et al. 1999, Grondelaers et al. 2001, Grondelaers et al. 2008, Speelman & Geeraerts 2009). It would take us too far afield to go into the journalistic practices and norms in Belgium and in the Netherlands, but it is not

unlikely that the Netherlandic newspapers are gravitating more quickly towards informal language use than the Belgian newspapers. It is possible that the intensifying fake reflexive resultative construction, being a subjective, expressive construction (cf. Chapter 2), is still slightly more associated with informal registers in Belgian Dutch, whereas this is not (or no longer) the case in present-day Netherlandic Dutch, where such register differentiation is much less prominent. This could explain why the construction is overall more frequent in Netherlandic Dutch. In that light, it is perhaps surprising that the construction actually appears to show more lexical variation and creativity in the intensifiers in Belgian Dutch, but the difference was admittedly rather small. At any rate, the results presented in this chapter provide clear evidence that the construction is by no means infrequent in the journalistic genres, and that journalists do appear to play around with the possibilities it offers, even (or especially so) in Belgian Dutch. Due to the lack of historical data for Belgian Dutch, we are unable to investigate whether and how these national differences have developed over time. There is some evidence in the literature that recent decennia have seen lexical convergence of standard Belgian and Netherlandic Dutch (Geeraerts et al. 1999, Daems et al. 2013, Zenner et al. 2013), but it is not entirely clear what this means for the construction under investigation. However, we did have sufficient diachronic data for Netherlandic Dutch, which makes it possible to track the development of the intensifying fake reflexive resultative construction over the past two centuries. This will be presented in the next chapter.

Chapter 5 Diachronic variation: the development of the intensifying fake reflexive resultative construction in Netherlandic Dutch between the early 19th and late 20th century

The previous chapter showed that the intensifying fake reflexive resultative construction in present-day Dutch allows for a great deal of lexical variation in its verb slot and its intensifier slot at the macro-level, while at the same time displaying signs of low-level pockets of productivity and conventional collocations. If we compare this to what we already know about the construction before the 19th Century, at which point its use was mostly limited to fixed expressions with *lachen* 'to laugh' and a small set of verb-intensifier combinations with *dood* 'dead' (see Ch2, §2.2.2.3), it appears that the construction has undergone important changes over the past two centuries. In this chapter, we will investigate how the construction has developed from the restricted set of attested uses in the early 19th Century to the present-day situation in which over 120 intensifier types enter into both conventional and productive combinations with more than 250 different verbs. While this chapter will primarily focus on changes within the Delpher data set, we will sometimes refer to the SoNaR data that were presented in the previous chapter. Only the first part of the subsections from Chapter 4, i.e. the "Synchronic use" sections, will be used as a point of comparison here: as we do not have historical data for Belgian Dutch, it will not be investigated how the variation between the two national varieties of Dutch has developed over time. The structure of this chapter is largely parallel to that of the previous chapter. We begin with a preliminary look into general frequencies and slot fillers in §5.1, before moving on to the interaction between the verb and intensifier slots in the collocational patterns in §5.2. Section 5.3 looks into frequency-based and semantic aspects of (historical) productivity on the basis of the models by Baayen and Barðdal (§5.3.1 and §5.3.2, respectively), focusing both on the macro-level and the intensifier-specific and verb-specific subschemas. These results will allow us to track shifts in the taxonomic organisation of the constructional network in

the final subsection (§5.4). Finally, §5.5 summarises the main findings of the diachronic part of the investigation and presents a short discussion on constructional changes in context.

5.1 A preliminary look into frequency development and changes in the slot fillers

The first subsection of Chapter 4 presented a detailed overview of the general frequency and use of the intensifying fake reflexive resultative construction in present-day Dutch. We studied the concrete instantiations of the different elements of the construction, paying attention to formal and semantic properties, but also to general frequency information. While historical linguists have often focused on the former (i.e. semantic or formal changes of a construction), it has recently come to their attention that *frequency* changes may also be informative in their own right (Hilpert & Gries 2009, Hilpert 2013, Diessel & Hilpert 2016). With the increase in digitally available diachronic corpora which can be controlled for certain parameters like time period or genre, multiple methods have been developed for interpreting the observed frequency developments in diachronic data, rather than just using frequency data as a way of descriptive illustration (Hilpert & Gries 2009). The present investigation of the diachronic changes in the intensifying fake reflexive resultative construction therefore begins with a discussion of the overall frequency development. Table 5.1 summarises the most important frequency measures.

Table 5.1. Frequency development of the intensifying fake reflexive resultative construction (type and token frequencies)

	1830s	1850s	1870s	1890s	1910s	1930s	1950s	1970s	1990s	SoNaR
Corpus size (in million words)	51.3	115.1	306.2	300.8	294.7	303.5	303.7	299.4	297.0	62.4
Absolute frequency	6	8	56	96	156	271	574	842	2,035	1,042
Normalised frequency	1.17	0.69	1.83	3.19	5.29	8.93	18.90	28.13	68.51	142.5
Verb types	5	6	17	25	38	39	53	96	131	137
Intensifier types	4	8	15	23	29	37	71	96	115	68

Eyeballing the frequencies suggests a substantial increase in the extent of use of the intensifying fake reflexive resultative construction. In the past hundred years, the construction has increased from just under 100 instances (normalised frequency of 3.19) to over 2,000 instances (normalised frequency of 68.51), which equals a multiplication by a factor of 21. This substantial increase is even more obvious if we visualise it in Figure 5.1. It appears that the construction first underwent a rather steady, gradual increase up until the 1930s, at which point the line starts to rise more steeply, leading up to an especially large frequency jump between the 1970s and the 1990s. According to the assumption that text frequencies reflect the familiarity of speakers with a specific construction at a given point in time (Hilpert 2017: 53), the frequency development could indicate that speakers of Dutch are becoming increasingly familiar with the intensifying fake reflexive resultative construction. Occasionally, we come across a metalinguistic comment that attests to the linguistic awareness of language users. One of these comments was found as early as 1903, suggesting that the intensifying construction had already reached a certain level of salience at that time.

(199) Zonder twijfel, de overdrijving zit bij ons in de lucht en in ‘t bloed. [...] De **een lacht zich ziek**, de andere **een bochel**, ja een derde **lacht zich dood**. (De Gooi- en Eemlander, 1903)
‘Without any doubt, exaggeration is in our nature. This one laughs “himself sick”, another one laughs “himself a hunch”, a third person laughs “himself dead”.’

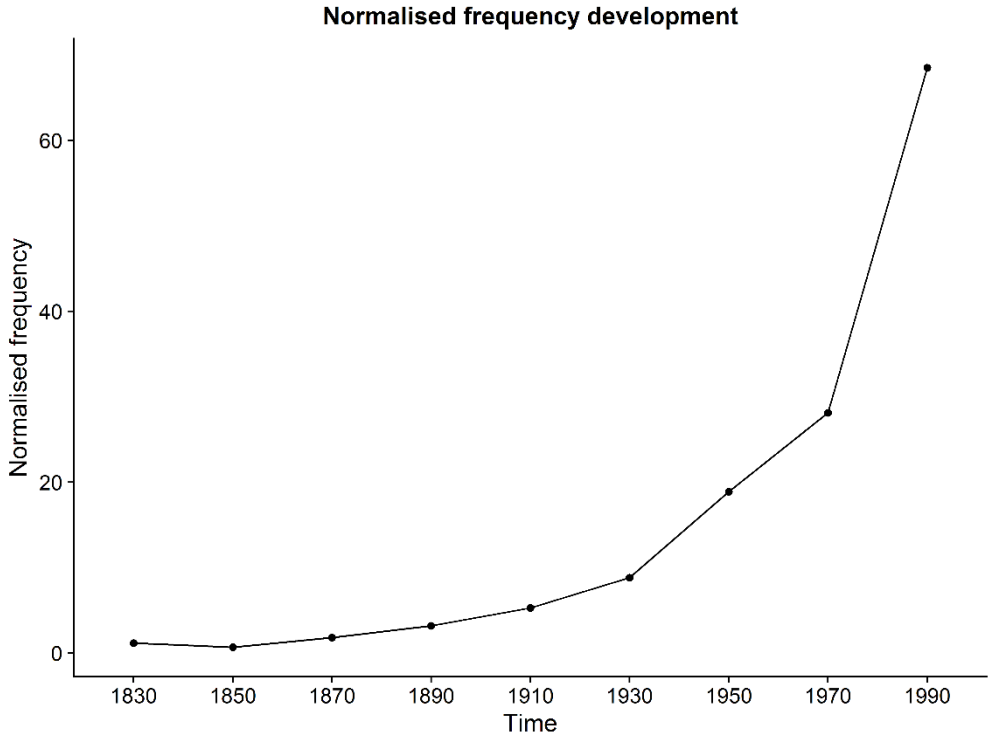


Figure 5.1. Frequency development of the intensifying fake reflexive resultative construction since the early 19th Century

A widespread statistic to test for the significance of a monotonic trend (viz. a simple upward or downward trend) is Kendall's tau. If we correlate the decades with the normalised frequencies, we get an overall correlation coefficient of 0.944 with a p-value of $p<.001$, indicating a highly significant upward trend. However, it would be interesting to describe this trend in more precise terms and explore whether there are any substructures in the data. This can be done using the variability-based neighbour clustering (VNC) analysis, an adaptation of hierarchical agglomerative clustering that takes into account the temporal sequentiality of the data (Gries & Hilpert 2008, 2012, Hilpert & Gries 2009). It is common practice to divide one's data into arbitrary equidistant time periods like decennia or half centuries. The VNC analysis, in contrast, is a bottom-up, data-driven approach to the periodisation of data that takes the underlying structure in the actual data as the basis to discern "phases" which are not necessarily equidistant. Without going into the details of the algorithm, for which we refer to the above references, the VNC analysis merges sequential time periods that are similar in terms of a specific quantitative criterion (e.g. token frequency) and sets them apart from time periods that behave rather differently. The analyst has to decide what counts as "similar" or "different" behaviour, that is, he/she has to decide which measure will be at the basis of the similarity calculations. In addition, he/she has to choose a similarity measure, i.e. how the similarities are calculated, and an amalgamation rule, i.e. how the merging is performed. In the simple case of one frequency observation per time period, the similarity measure is often based on the pairwise distance between data points (e.g.

standard deviation) and the default amalgamation method averages over the individual values of the time periods.

The VNC analysis is best applied to data sets that are truly diachronically continuous, i.e. without any missing periods. Our data is not ideal in that regard because we have left out the even-numbered decennia (‘20s, ‘40s, ‘60s and ‘80s), but as we do have all consecutive odd-numbered decennia (‘10s, ‘30s, ‘50s, ‘70s and ‘90s), we applied the VNC analysis to our token frequency data, see Figure 5.2.

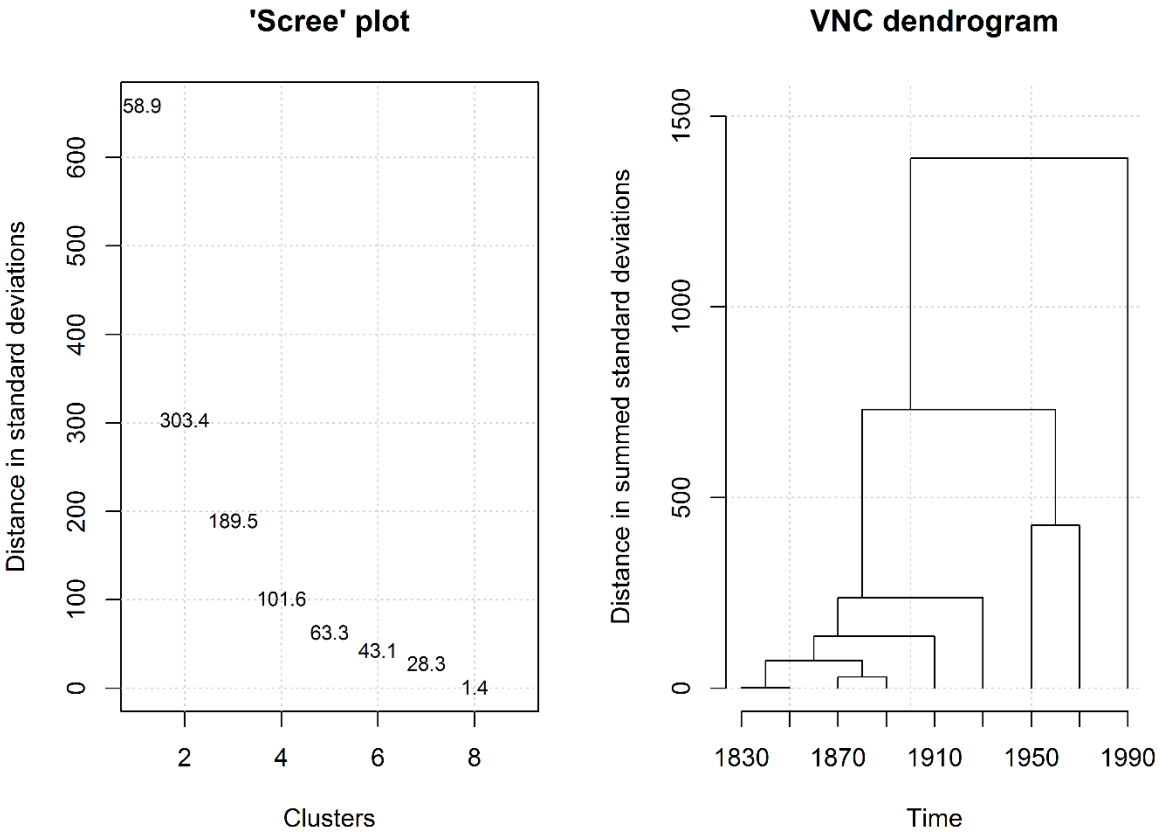


Figure 5.2. VNC scree plot and dendrogram for absolute token frequency

The scree plot on the left gives an indication of how many clusters should be discerned within the diachronic development: the ideal number of clusters is reached when the difference between the points on the scree plot is no longer substantial. In this case, we observe a rather steep decline until the point of four clusters, at which point it levels off horizontally. Based on the shape of the dendrogram in the right panel then, we can determine what these four clusters are. It appears that the first cluster ranges from the 1830s until the 1910s, the second cluster consists only of the 1930s, the third cluster groups the 1950s and the 1970s and the fourth cluster is limited to the 1990s. On the basis of the frequency graph, it was already observed that the 1930s appear to have been some kind of breakthrough point at which the construction started to gain in frequency at a more rapid rate. Remarkably, the scree plot shows an especially large gap between the

first and the second point, suggesting that at the coarsest level of a two-cluster option, there is an enormous difference between the 1990s (i.e. the cluster that is added last) and all previous periods. This result also aligns with the frequency jump that we observed between the 1970s and 1990s in Figure 5.1.

The VNC analysis is generally calculated on the basis of one frequency measure (often raw token frequency, see Hilpert 2012, Lorenz 2012, Onysko & Calude 2014, relative frequency, see Hilpert 2011, Rosemeyer 2014, or collostruction strength, cf. Gries & Hilpert 2008, though see Hilpert 2013 for a periodisation according to several measures simultaneously). Depending on the measure that is selected as input for the VNC analysis, the analysis may suggest a different periodisation as outcome (see Perek & Hilpert 2017, who show that a periodisation on the basis of distributional semantic information differs from purely quantitative measures, such as token or type frequency). That is, it is possible that other frequency measures do not follow the exact same trend as the overall frequency of occurrence, in which case the VNC analysis would reveal such differences in the underlying data structure. By that same logic, it may also reveal similarities in the development of multiple quantitative measures. We can test this by using type frequency instead of token frequency as input. Figure 5.3 shows the overall frequency development of the intensifier types and the verb types.

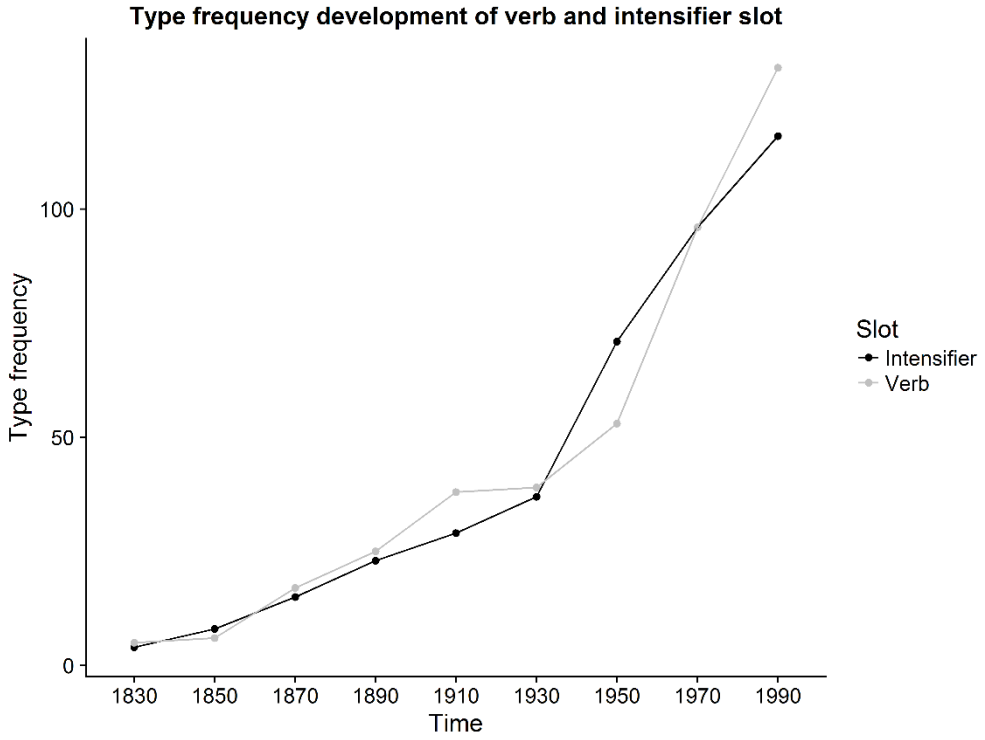


Figure 5.3. Frequency development of the different verb and intensifier types in the intensifying fake reflexive resultative construction since the early 19th Century

Much like the overall frequency of the construction, both the numbers of attested intensifier and verb types have increased substantially since the early 19th Century. Kendall's tau confirms that there is a significant upward trend in both the intensifier and

verb slots, with τ=0.94 and τ=0.97, p<.001 respectively. Such a development is reminiscent of the concept of diffusional change in De Smet (2013). Diffusion covers a variety of more specific changes which are generally conceived of as "the gradual unidirectional expansion of a linguistic item over a new range of lexicogrammatical concepts" (2013: 45, see 46-61 for an overview of different subtypes of diffusional changes). Their gradualness lies in the fact that the changes do not happen all at once and often proceed at a different pace in different contexts. They are unidirectional in that the changes have to be consistent, "without any serious fallback or fluctuation" (2013: 3): once a certain change has happened, it is generally not reverted. Consistent collocational expansion to new verbs, as shown in Figure 5.3, is one of the changes that qualify as diffusional. Interestingly, both curves start out at around the same point and, in spite of some small variations in the curvature, they continue to run largely parallel for the entire period under investigation. This seems to suggest that the construction is expanding on multiple fronts by concurrently coining new intensifiers as well as extending (existing) intensifiers to previously unattested verb types. The curve that captures the expansion of the repertoire of intensifier types shows a distinct kink around the 1930s. Up until the 1930s, the intensifier slot appears to slowly but surely expand its range, gaining 5 to 10 new types in each successive decennium. Between the 1930s and the 1950s, the intensifier slot suddenly jumps from 37 to 71 different intensifier types, nearly doubling in frequency. After the 1950s, the slope continues to increase slightly less steeply, although we still get another increase of 25 types between the 1950s and the 1970s and of 19 more types between the 1970s and the 1990s.

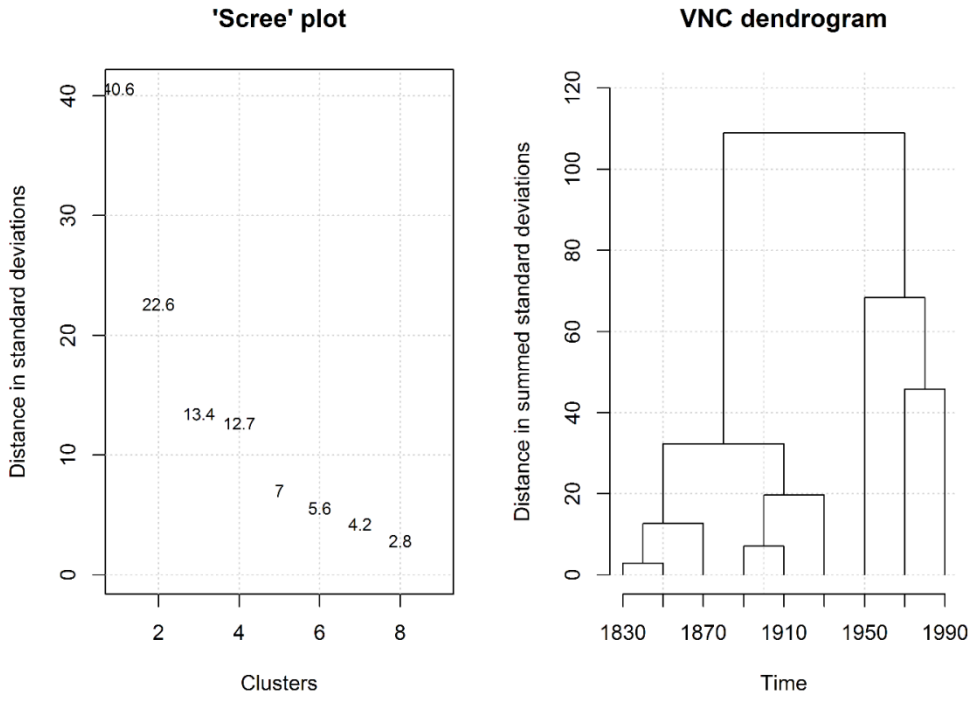


Figure 5.4. VNC scree plot and dendrogram for intensifier type frequency

Judging by the scree plot, we could distinguish either three or five clusters. Based on the shape of the dendrogram, the solution of three clusters presents itself as somewhat more elegant and it also aligns with the earlier observations on the basis of the frequency graph. We get a first cluster that covers the entire century from the 1830s until the 1930s, a second cluster featuring only the 1950s and a third cluster with the 1970s and the 1990s. The fact that the 1950s now form a separate cluster in the type frequency VNC analysis is explained by the sudden increase in intensifier types between the 1930s and the 1950s.

For the verb slot, there is also a clear overall upward trend in Figure 5.3, but it is slightly less smooth. From the 1850s onwards, the construction gradually starts to expand to new verb types, but this process is temporarily halted or slowed down between the 1910s and the 1930s. While the type explosion of the intensifier slot took place between the 1930s and 1950s, the key phase for the verb slot appears to be between the 1950s and the 1970s, going from 53 to 96 different verb types.

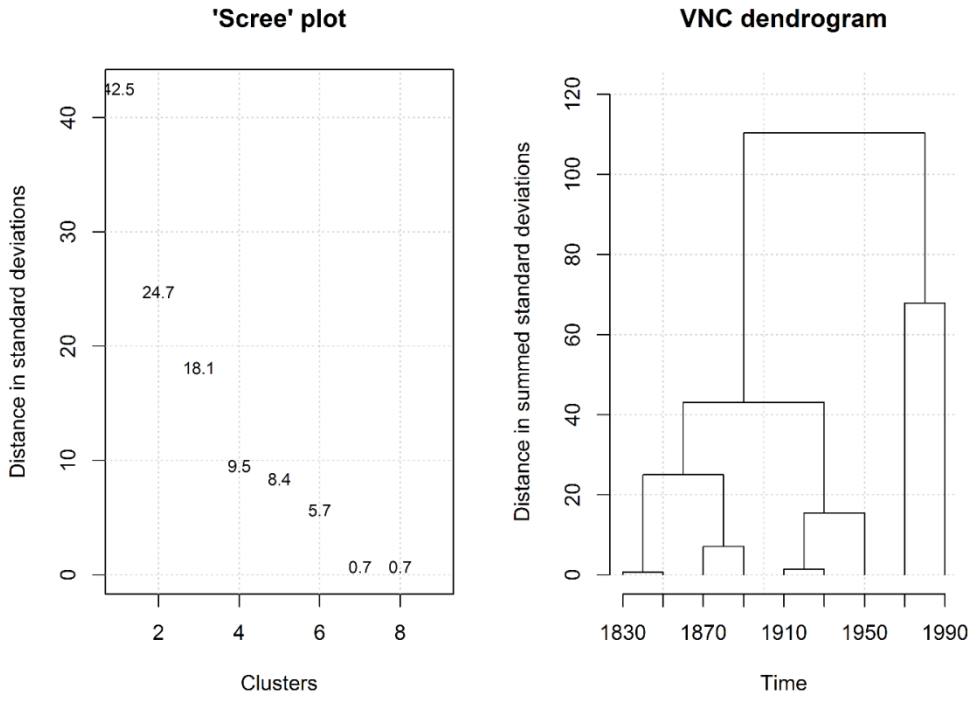


Figure 5.5. VNC scree plot and dendrogram for verb type frequency

The scree plot in Figure 5.5 suggests a division into four time phases (or seven, but in that case we would have too many clusters consisting of just one decennium), in which we recognise the phases from the frequency development in Figure 5.3. The first phase comprises the 1830s and the 1850s, the second phase groups together the 1870s and the 1890s, the third phase ranges from the 1910s to the 1950s and finally the 1970s and 1990s are merged in the fourth phase.

To summarise, whereas Kendall's tau can tell us that the intensifying fake reflexive resultative construction has increased its extent of use in terms of token frequency and verb and intensifier type frequencies, the frequency development graphs and the VNC

analyses reveal that these increases have progressed in slightly different ways. Still, all graphs that we have analysed so far hint at important changes around the middle of the 20th Century. From that moment onwards, we observe an acceleration of the overall increase in frequency and there appear to have been some crucial changes in the instantiation of the verb and intensifier slots. As we aim to provide a detailed account of the (recent) changes in the intensifying fake reflexive resultative construction, we have decided to also include the intervening decennia from the 1930s onwards.[54] In total, we now have 6,135 instances of the construction, i.e. the original 5,325 tokens plus the instances from the newly included decades. A new overview of the most important frequency information, with the 1940s, the 1960s and the 1980s included, can be found in Table 5.2.

Table 5.2. Frequency development of the intensifying fake reflexive resultative construction (type and token frequencies), plus 1940s, 1960s and 1980s

	1830s	1850s	1870s	1890s	1910s	1930s	1940s	1950s	1960s	1970s	1980s	1990s
Corpus size (in mill words)	51.3	115.1	306.2	300.8	294.7	303.5	295.4	303.7	304.9	299.4	296.7	297.0
Absolute frequency	6	8	56	96	156	271	316	574	498	842	1,277	2,035
Normalised frequency	1.17	0.69	1.83	3.19	5.29	8.93	10.73	18.90	16.33	28.13	43.04	68.51
Verb types	5	6	17	25	38	39	42	53	56	96	94	131
Intensifier types	4	8	15	23	29	37	56	71	65	96	76	115

Figure 5.6 only zooms in on the period from the 1930s until the 1990s. In the left panel, we see the normalised frequency development, the right panel gives the type frequency development for verbs and intensifiers.

[54] Concretely, the search query from round one in Ch3, §3.2.2.1 yielded 28,656 hits for the 1940s, 31,152 for the 1960s and 32,319 for the 1980s. If any new intensifiers (i.e. intensifiers that were not part of the original set of 171 intensifiers retrieved in round one, cf. Ch3, §3.2.2.1) were retrieved during this step, we would have to launch an additional round two query (i.e. [REFL INT]) for these new intensifiers in *all* decennia. As this did not turn out to be the case, the data sets for the other decennia remained unchanged. We proceeded to step two with the original set of 171 intensifiers, which yielded another 12,404 results for the 1940s, 11,672 for the 1960s and 13,952 for the 1980s.

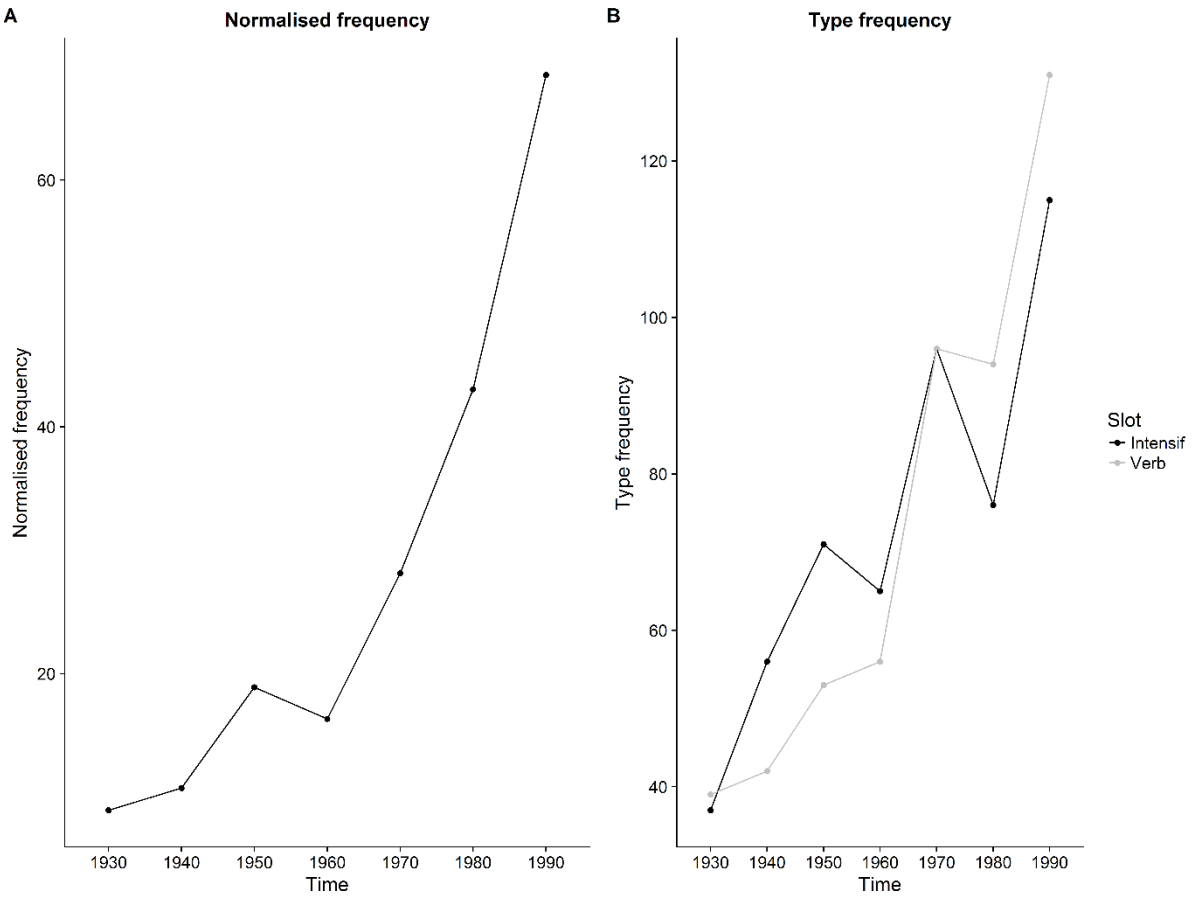


Figure 5.6. Frequency development from the 1930s until the 1990s

Curiously, the inclusion of the intervening decennia has added striking fluctuations to the frequency development curves. The intensifier type frequency curve is especially interesting in that it shows a marked rise-and-fall pattern, with peaks in the 1950s and 1970s. The curvature is also somewhat irregular for the verb types: the 1950s and the 1970s also see a hump in the curve, but there is no relapse in the 1960s (on the contrary, there is a small increase by 3 verb types) and between the 1970s and 1980s, the construction only loses 2 verb types (compared to 20 intensifier types). In that light, it needs to be pointed out that extension to new verb types and extension to (i.e. creation of) new intensifiers types are quite different in nature. Concretely, once a new verb type has been recruited to the construction, it is very unlikely for it to disappear again. Individual verbs are not likely to be removed from the pool of potential (though not necessarily attested) verb slot fillers, unless the verb as such disappears from the language. While it is also true that an intensifier that is unattested is not necessarily impossible or inexistent (cf. Ch4, §4.2.1 on the difference between unattested and ungrammatical), we will see that some intensifiers that were once used somewhat frequently in the construction do seem to have disappeared from the repertoire in a later stage. As was argued in Chapter 2, §2.3.2, the domain of intensification is generally characterised by renewal and innovation: intensifiers may receive competition from new, creative and more expressive alternatives and eventually cease to be used

completely. At first blush, this rise-and-fall structure in Figure 5.6 does not tally well with the idea of diffusional change, which is supposed to be characterised by a unidirectional trend rather than the kinds of fluctuations that are observed here (cf. supra). Still, we do see that the brief setback ultimately does not prevent the overall expansion of the verb (or intensifier) slot, so the upward trend is still "unidirectional", it is just not exactly smooth or linear. One possible interpretation of the fluctuations is that the 1950s and the 1970s were extremely prolific in spawning new intensifiers, for whatever reason, which might have triggered a slight bump in verb frequency as well. Judging from the dips in the 1960s and the 1980s, many of these newly created intensifiers did not survive into the next decennium (see §5.1.3.2). Interestingly, although the type frequency peak of the 1950s is mirrored by a peak in the normalised frequency curve, the same cannot be said for the 1970s. While there was a decrease in intensifier and verb type frequencies between the 1970s and the 1980s, the overall frequency of the intensifying construction continued to follow the general upward trend.

As we now do have a truly diachronically continuous period in our data set, we can check for additional substructures in the data by reapplying the VNC analysis to this subperiod. Figure 5.7 gives the results for the absolute frequency development between 1930 and 1995.

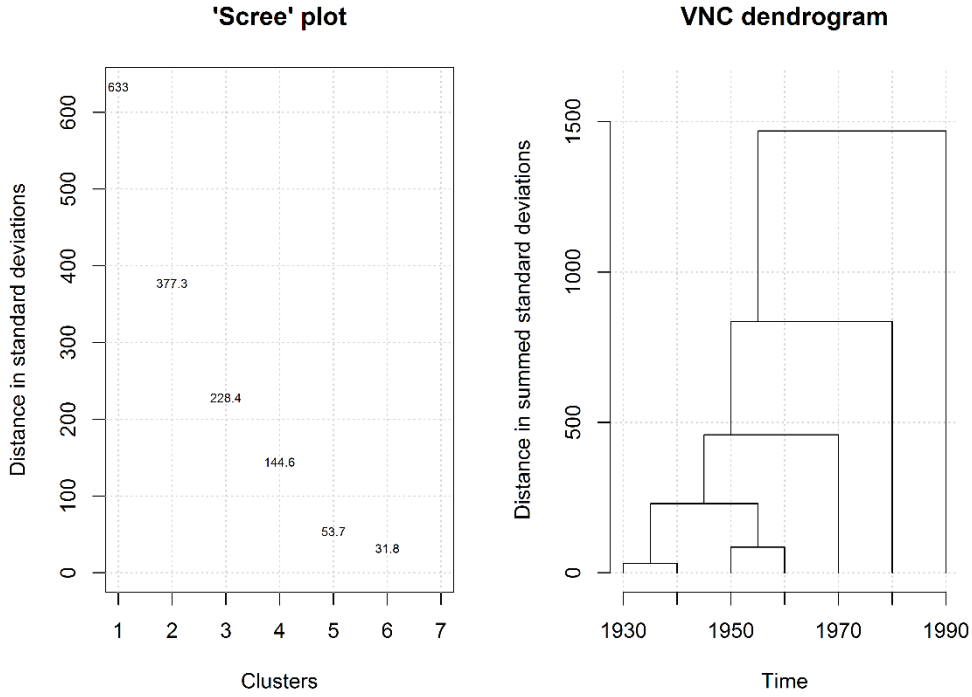


Figure 5.7. VNC scree plot and dendrogram for absolute token frequency (1930-1995)

The scree plot in the left panel only levels off at six clusters, at which point most clusters consist of only one time period. While it can be meaningful to single out one or two decennia to signal their uniqueness with respect to the surrounding decades, there is no

point in using a cluster analysis if *most* of the clusters are individual time periods. What the analysis tells us is that almost each consecutive decennium from the 1930s onwards shows a substantial change in frequency compared to the previous decennium.

In the following paragraphs, we will zoom in on the different elements of the intensifying fake reflexive resultative construction. A thorough analysis of the lexical items that were added to and/or removed from the repertoire of potential slot fillers may also offer a possible explanation for the fluctuation that was observed in the type frequency curves in Figure 5.6.

5.1.1 Verb

The Delphcorp data set of 6,135 instances of the intensifying fake reflexive resultative construction features 245 different verbs.

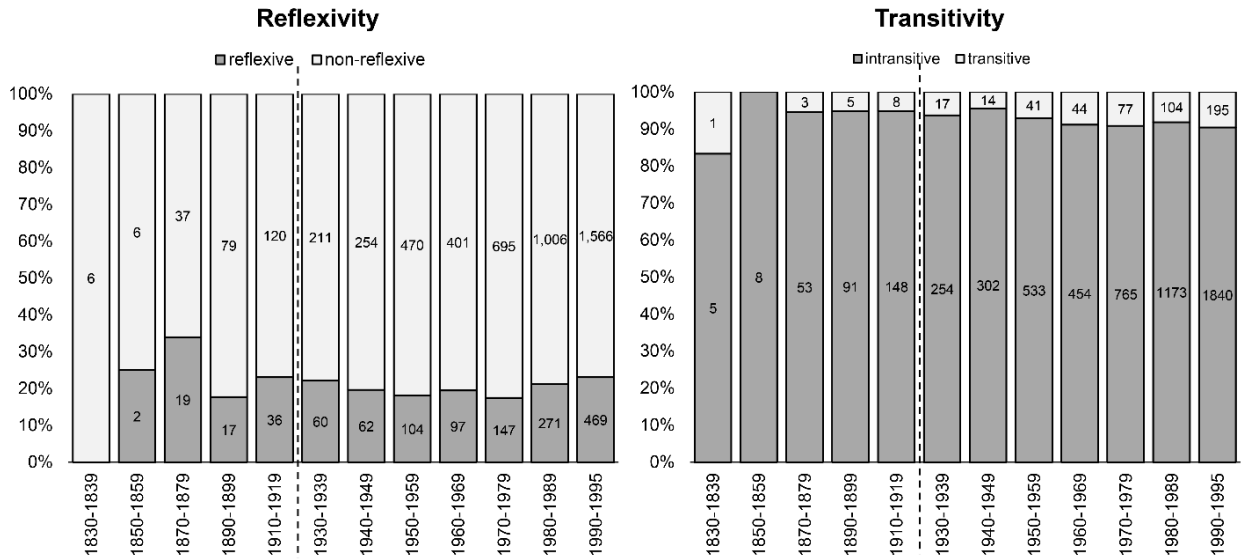


Figure 5.8. Development of the proportions of (non-)reflexive and (in)transitive verbs in Delphcorp[55]

If we look at the proportions of reflexive versus non-reflexive and transitive versus intransitive verbs in Figure 5.8, we see that the dominance of non-reflexive intransitive verbs that was attested for present-day Dutch in Chapter 4 is also historically robust. Like in present-day Dutch, the inherently reflexive verbs have a more important role in the construction than one would guess by looking at Figure 5.8. Given that there are only 7 inherently reflexive verbs in the entire data set and 4 of these are relatively infrequent, it is remarkable that the 3 remaining reflexive verbs, viz. *zich ergeren* 'to be annoyed', *zich schamen* 'to be embarrassed' and *zich vervelen* 'to be bored', consistently account for about

[55] The dashed lines between the 1910s and 1930s indicate that, from that point onwards, the even-numbered decennia (i.e. the 1940s, 1960s and 1980s) were included in the analysis.

a quarter of all data in each period. Their historical importance is confirmed by Table 5.3, which gives the top ten verbs for three decennia in the corpus, representing the second half of the 19th Century (1870-1879), the first half of the 20th Century (1930-1939) and the late 20th Century (1990-1995) (a full overview of all verbs per decennium can be found in Appendix V-1). Clearly, the reflexive verbs have played an important part in the overall frequency development of the construction. As was argued in Chapter 4, the strong representation of the reflexive verbs is quite likely based on their semantics, rather than their inherent reflexivity. They express human experiences that are particularly prone to intensification, viz. annoyance, embarrassment, boredom, much like other top verbs like *lachen* 'to laugh' and *schrikken* 'to be startled'.

Table 5.3. Evolution of the top 10 verbs in the 1870s, 1930s and 1990s

	1870-1879	1930-1939	1990-1995
1.	**zich ergeren (15)**	**lachen (64)**	schrikken (470)
	'to be annoyed'	'to laugh'	'to be startled'
2.	**lachen (13)**	schrikken (47)	**zich ergeren (270)**
	'to laugh'	'to be startled'	'to be annoyed'
3.	**werken (7)**	**zich vervelen (20)**	**werken (224)**
	'to work'	'to be bored'	'to work'
4.	denken (3)	**werken (19)**	**lachen (166)**
	'to think'	'to work'	'to laugh'
5.	kniezen (2)	**zich ergeren (18)**	**lopen (134)**
	'to mope'	'to be annoyed'	'to run'
6.	**lopen (2)**	**zich schamen (18)**	**zich schamen (106)**
	'to run'	'to be embarrassed'	'to be embarrassed'
7.	**zich schamen (2)**	denken (12)	**zich vervelen (85)**
	'to be embarrassed'	'to think'	'to be bored'
8.	**zich vervelen (2)**	**lopen (8)**	piekeren (60)
	'to be bored'	'to run'	'to worry'
9.	zoeken (2)	zoeken (7)	trainen (45)
	'to search'	'to search'	'to train'
10.	blazen (1)	kniezen (6)	betalen (37)
	'to blow'	'to mope'	'to pay'

Table 5.3 further shows that there is remarkable continuity in the verbs that occupy the top positions: six verbs, indicated in bold, are featured in the top ten of all three decades. Perhaps surprising is the absence in the top ten of the 1870s of the verb *schrikken* 'to be startled', given that it is currently the most frequently intensified verb by a large margin. In fact, *schrikken* 'to be startled' is not attested at all in the construction before the 1910s.

While most verbs have increased in absolute frequency, not all verbs show exactly the same development: the increase is much more substantial for some verbs than for others. That is even more obvious if we look at Figure 5.9, which compares the frequency developments of the top twenty verbs (based on the summed frequencies over the entire data set) from the 1930s onwards.

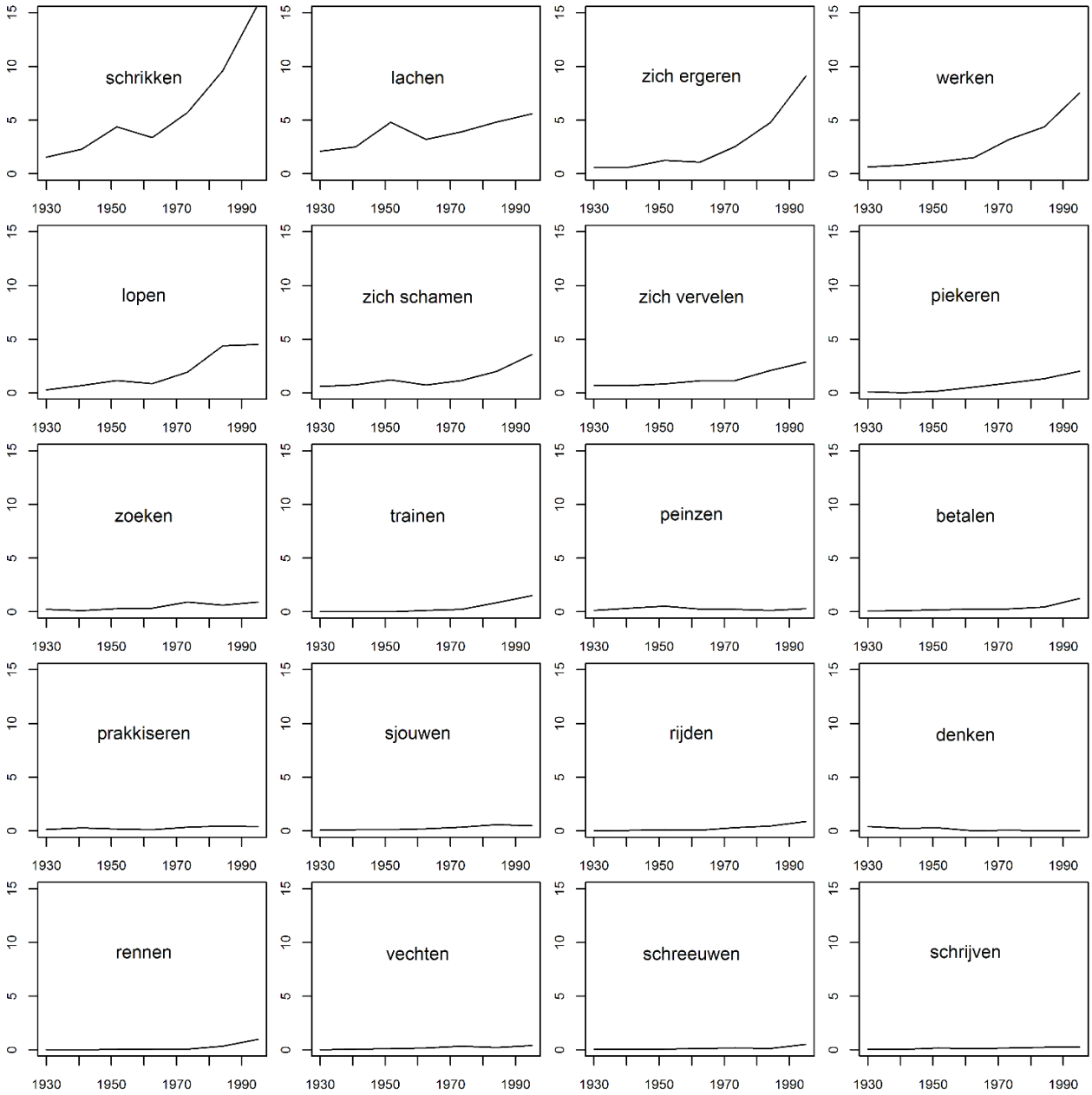


Figure 5.9. Normalised token frequency developments of the 20 most frequent verbs (1930s-1990s)

The lower panels are generally characterised by stability across the entire period and there are no examples of clear declining trajectories. Only the cognitive verbs *denken* 'to think' and *peinzen* 'to ponder' do appear to have decreased in frequency, albeit slightly. In the top panels, there are some examples of verbs that have achieved some moderate success only recently, e.g. *piekeren* 'to worry', *trainen* 'to train', *rennen* 'to run' and *betalen* 'to pay'. The verb *lopen* 'to run' appears to be levelling off in the most recent decennia. The top panel rows suggest that the overall frequency increase of the construction was carried by a number of highly frequent verbs which started to increase exponentially in the second half of the 20th Century, especially *schrikken* 'to be startled', *zich ergeren* 'to be annoyed' and *werken* 'to work'. The verb *lachen* 'to laugh', which is one of the earliest

experience verbs to be used in the construction, shows a more gradual frequency increase. Interestingly, both the verbs *lachen* and *schrikken* have a noticeable peak in the 1950s, which mirrors the peak in overall frequency of the construction that was observed in Figure 5.6.

Following Hilpert (2015a), we can measure the importance of these top twenty verbs by setting them off against the remaining 225 verbs. The verbs were arranged in such a way that the upper area of Figure 5.10 fades from black to dark grey in descending frequency; the lightest grey area at the bottom of the graph represents all other 225 verb types.

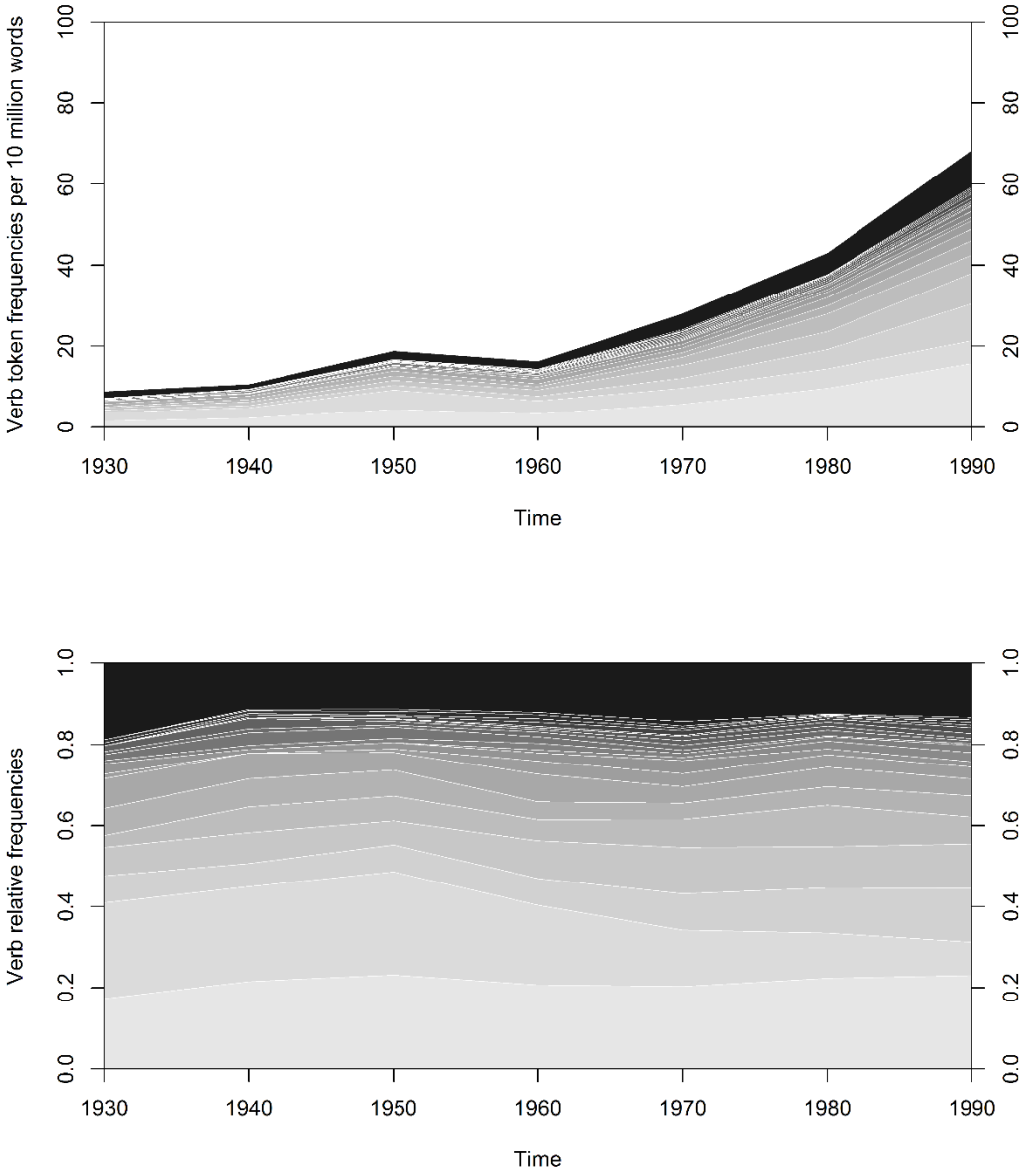


Figure 5.10. Frequency developments of 20 frequent verb types versus 225 infrequent verb types. (upper panel: normalised frequency development, lower panel: relative frequency development)

Figure 5.10 shows that the top twenty verbs have consistently represented around 80% of the data set. This entails that the noticeable increase in verb type frequency and hapax count must be situated in the remaining 20% of the data. If we look at the lower frequency verbs that are recruited to the verb slot, we see that the classes of emotional and cognitive experience verbs and the verbs denoting physical activities, both of which were already well-populated in the early decades, continue to attract new members (see the non-exhaustive lists in (i) and (ii) below). These semantic domains have become more dense and are still very prominent in present-day Dutch.

(i) Experience/cognitive verbs: *filosoferen* 'to philosophise', *giechelen* 'to giggle', *mijmeren* 'to muse', *mediteren* 'to meditate', *schreien* 'to weep', *tobben* 'to fret'...
(ii) Physical activity verbs: *draven* 'to trot', *fietsen* 'to cycle', *rijden* 'to ride', *schaatsen* 'to skate', *spelen* 'to play', *zwemmen* 'to swim', *zwoegen* 'to labour'

At the same time, there are indications that the verb slot is widening its semantic scope. The other semantic categories that were established in Chapter 4 on the basis of the synchronic data, have not all been represented since the very beginning. Only in the 1890s do we start finding multiple examples of verbs of noise emission, e.g. *schreeuwen* 'to scream', *snikken* 'to sob' (which could also be an experience verb), *zingen* 'to sing', see (200).

(200) En 't kind **schreeuwt zich een ongeluk**, ik begrijp er niets van; niets helpt. (Delphcorp, 1890-1899)
and the child screams itself an accident [...]
'And the child is screaming from the top of its lungs, I don't understand. Nothing helps.'

The verbs of consumption *drinken* 'to drink' and *eten* 'to eat' were already used quite early on, but the verbs that denote smoking or taking drugs are of a more recent date and relatively infrequent overall. There are also some early attestations of verbs of communication, see (201), but it is only in the second half of the 20th Century that we start observing some more lexical variation in this category (cf. also the verbs of new instruments of communication like *mailen* 'to email' or *sms'en* 'to text' in SoNaR-NL, Chapter 4).

(201) Ze loosen hun diepe zuchten en roerende klachten in ingezonden stukken, **argumenteeren** en molenwieken **zich half dood**. (Delphcorp, 1910-1919)
[...] argue and wing themselves half dead
'They heave a deep sigh and express touching complaints in letters to the editor, they argue and stir like no other.'

In the second half of the 20th Century, we start coming across a variety of verbs that are not easily classified into one of the larger categories, denoting all kinds of concrete or abstract activities that generally do not require much physical effort, e.g. *investeren* 'to invest', *prijzen* 'to prize', *protesteren* 'to protest', *schakelen* 'to switch gears', or other verbs

that are hard to pin down to a specific category. Some examples are found in (202) to (204). Most of the new types that are added in the second half of the 20th Century belong to this "other" category (or "general activity" category, if you will).

(202) Ik zie daar een goed stuk brood en mijn vrouw **breit zich het apezuur**. (Delphcorp, 1980-1989)
[...] my wife knits herself the monkey-acid
'I spot a good loaf of bread and my wife is knitting her fingers off.'

(203) In de supermarkt **prijsden** de vakkenvullers **zich een ongeluk** om de inflatie (bijna 5000 procent in 1989) bij te houden. (Delphcorp, 1990-1995)
in the supermarket prized the shelf stackers themselves an accident [...]
'In the supermarket, the shelf stackers had to keep changing the prices to keep up with the inflation (nearly 5,000 percent in 1989).'

(204) Een drumcomputer **samplet zich suf** - dat hij was uitgenodigd mag als een bewijs van de democratische inslag van de trommels van de planeet worden beschouwd. (Delphcorp, 1990-1995)
a drum-computer samples itself drowsy [...]
'A drum computer samples hit after hit – the fact that he was invited is proof of the democratic streak of the drums of this planet.'

In sum, we do see some evidence of semantic broadening beyond the frequent emotive or cognitive experience verbs and the verbs of physical activity, which suggests that the verb slot has increased its schematicity over time. However, given the low frequency of the construction in the early stages, we cannot say with certainty whether this is because these kinds of verbs were really *impossible* in the construction or because we just did not come across any attestations.

5.1.2 Reflexive Pronoun

In Chapter 3, we discussed a number of changes in journalese, i.e. the language used in journalistic genres, which were brought together under the terms "informalisation" or "conversationalisation". One of the studies that were mentioned was the comparative corpus study of Vis (2011), in which Dutch newspapers from 1950 were compared to newspapers from 2002. Although there was an obvious style shift between the mid-20th Century and the 21st Century, it was not so much the case that the journalist himself adopted a different writing style. Rather, the main difference is that the 21st Century journalist is no longer the only voice of authority in his articles, insofar as he occasionally gives the floor to other news sources via reported discourse. In other words, what appears to be a style shift in newspaper language is actually an increase in the proportion of direct speech in news articles. In light of the subjective and expressive meaning aspect of the intensifying fake reflexive resultative construction (cf. Chapter 2), this is the type of construction we could expect to show up in direct speech. It is plausible, then, that

this general change in the newspaper genre explains at least part of the expansion that we saw earlier. Given the enormous amount of data that was processed manually (cf. Chapter 3) and the fact that such genre changes are not part of our research aims (see §5.5.2 for discussion), the instances in the data set were not annotated for whether they occurred in direct speech or not. In any case, as direct speech is generally characterised by first and second person expressions (cf. Ch3, §3.1.1), the changes in direct speech should only, or primarily, affect the instances of the construction featuring first and second person reflexive pronouns. Moreover, as not all first and second person examples are necessarily found in direct speech, direct speech examples only make up a fraction of the entire data set. That is, even though the increase in reported discourse may have contributed to the expansion of the construction, the impact is expected to be rather limited. We saw in Chapter 4 that a majority of 64% of the examples in the present-day data set feature third person pronouns, 27% contain first person pronouns and 9% second person pronouns. Figure 5.11 gives the relative frequency development of the different reflexive pronouns over time in the Delphcorp data set.

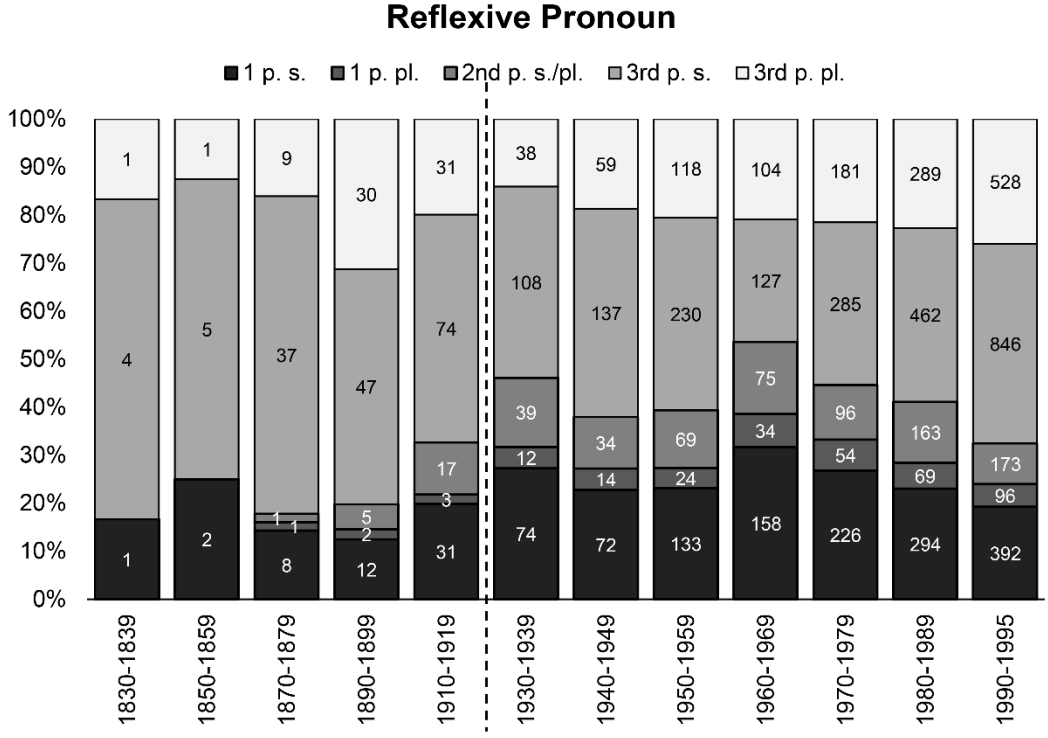


Figure 5.11. Proportion of reflexive pronouns in Delphcorp

Overall, the graph does not reveal any noteworthy trends, aside from the fact that the second person pronouns became slightly more frequent around the turn of the century. If there is an increase in reported discourse in newspapers overall, the impact on the use of the intensifying fake reflexive resultative construction is not obvious from the proportions of first and second person pronouns. If anything, there has been a general decrease in first and second person uses of the construction in the most recent decades. From the late 19th Century onwards, all types of reflexive pronouns have been fairly well-

represented in all time periods, even though the third person pronouns have rather consistently accounted for a (small) majority of the data. That is, the intensifying fake reflexive resultative construction does not appear to be confined to specific discourse contexts and can be used as a pattern of intensification when relating both one's own experiences and things that happened to someone else.

5.1.3 Intensifier

The entire Delphcorp data set contains 210 different intensifiers. To give a first idea of how the use of the intensifiers has changed over time, Table 5.4 lists the top ten intensifiers for the 1870s, the 1930s and the 1990s.

Table 5.4. Evolution of the top 10 intensifiers in the 1870s, 1930s and 1990s

	1870-1879	1930-1939	1990-1995
1.	**dood (26)**	**dood (111)**	rot (257)
	'dead'	'dead'	'rotten'
2.	ziek (8)	een ongeluk (28)	**dood (243)**
	'sick'	'an accident'	'dead'
3.	halfdood (4)	**suf (24)**	**suf (170)**
	'half-dead'	'drowsy'	'drowsy'
4.	**suf (4)**	halfdood (15)	kapot (137)
	'drowsy'	'half-dead'	'broken'
5.	de vingers krom (2)	krom (13)	te pletter (136)
	'the fingers bent'	'bent'	'to smithereens'
6.	krom (2)	slap (10)	groen en geel (107)
	'bent'	'weak'	'green and yellow'
7.	schor (2)	een aap (8)	**het vuur uit de sloffen (76)**
	'hoarse'	'a monkey'	'the fire out of the slippers'
8.	blind (2)	de ogen uit het hoofd (7)	uit de naad (71)
	'blind'	'the eyes out of the head'	'out of the seam'
9.	**het vuur uit de sloffen (1)**	**het vuur uit de sloffen (7)**	een hoedje (66)
	'the fire out of the slippers'	'the fire out of the slippers'	'a little hat'
10.	de ogen uit het hoofd (1)	schor (6)	wezenloos (66)
	'the eyes out of the head'	'hoarse'	'vacant'

Compared to the verb slot, which showed remarkably little diachronic variation in the top ten, there are some clear shifts in the intensifiers that hold the top positions. While there is some continuity between the 1870s and the 1930s, which also share the intensifiers *krom*, *schor*, *halfdood* and *de ogen uit het hoofd* in their top ten, there is little overlap between the 1930s and the 1990s. It appears that the 1930s mark the beginning of a period of important changes in the intensifier slot, and, as we will show below, many of the intensifiers that are in the top ten in present-day Dutch were in fact not yet (or barely) used in the 1930s. If we compare the 1990s to SoNaR-NL, then, it appears that the top ten has not changed a lot since the late 20th Century.

Figure 5.12 visualises the frequency development of absolute token frequencies, type frequencies and hapax counts for all syntactic categories that are represented in the intensifier slot. There are a number of important observations to be made. First of all, it can be gleaned from the general shape of the graphs in panels A and B that all major syntactic categories show an overall upward trend with respect to both token and type frequencies. The hapax graph in panel C has a more irregular structure, but for most categories the number of hapaxes is higher in the late 20th Century compared to the early 19th Century. Interestingly, the distinct rise-and-fall pattern that we observed for the type frequency of the intensifiers in general (see Figure 5.6) is reconfirmed for almost all syntactic categories. It was mentioned above that the 1950s and the 1970s were characterised by an unusual spike of creativity when it comes to introducing hapax intensifiers. It appears that the categories of nominal and NP+PP intensifiers were primarily responsible for the sudden increase in the 1950s; for the adjectival and prepositional intensifiers, we observe a decrease in panels B and C instead. For the 1970s, almost all categories show a peak in both type frequency and hapax count. It also seems that the overall increase in token frequency of the construction is spearheaded by the adjectival intensifiers, which exhibit an exponential growth in the second half of the 20th Century. The development of the type frequency presents a somewhat different view. From the beginning, the category of nominal intensifiers has managed to keep pace with the adjectival category with respect to its type frequency development. In the 1950s, the nominal category temporarily overtakes the adjectival category and continues to be close on its heels during the following decades, eventually coming out on top again by the 1990s (and in present-day Dutch, cf. Chapter 4). Moreover, the nominal category presents itself as the most productive category as of the mid-20th Century, in that it is responsible for a large proportion of one-off, creative intensifiers. This development is especially interesting because the nominal category has undergone a much less substantial token frequency increase compared to the adjectival category.

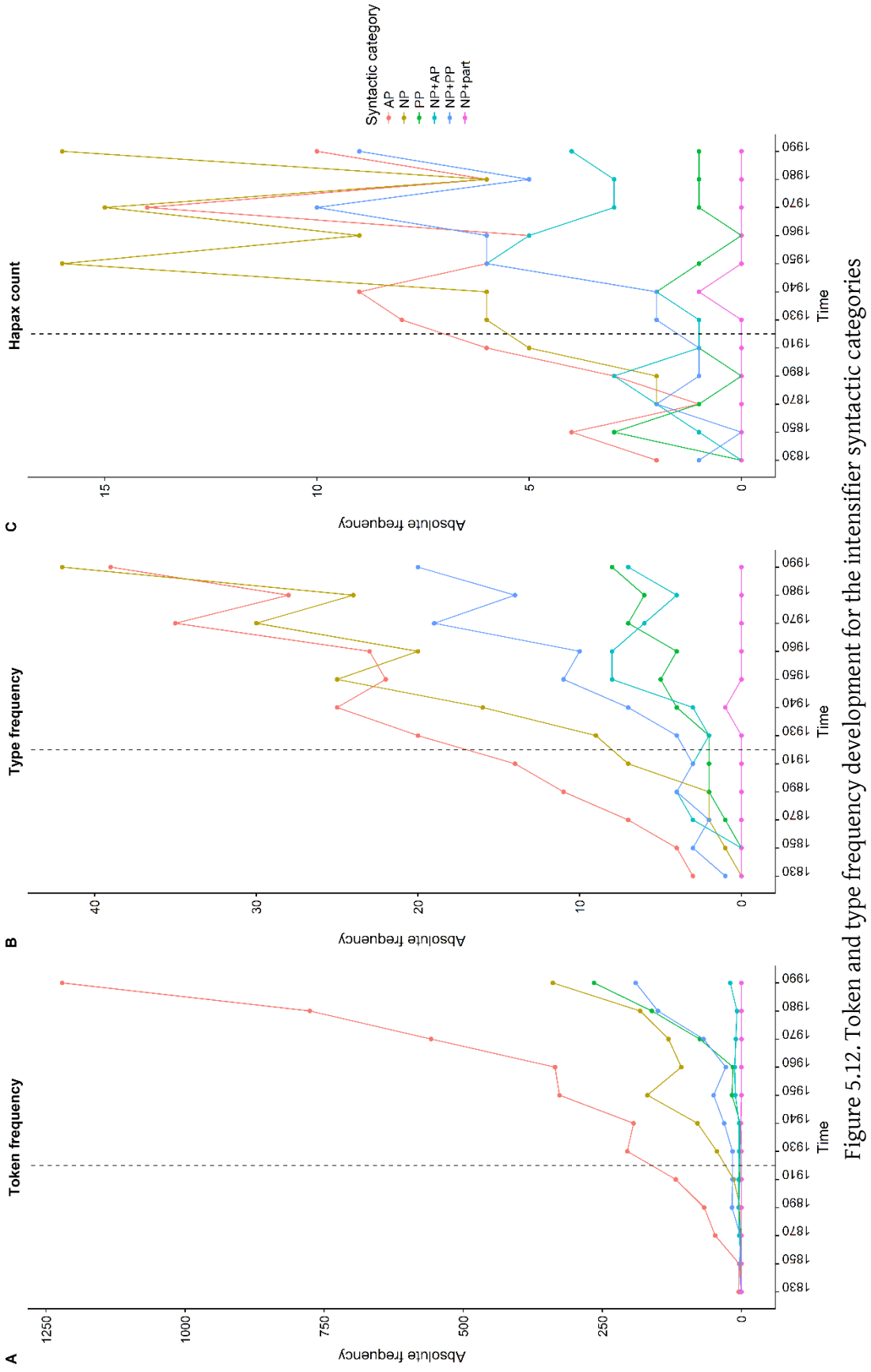


Figure 5.12. Token and type frequency development for the intensifier syntactic categories

If we look at the development of the degree of lexical variation, measured by the type-token ratio [TTR], of all syntactic categories in Table 5.5, we see that the adjectival category has pretty low values across the board.

Table 5.5. Development of the type-token ratios for the intensifier syntactic categories

	1830s	1850s	1870s	1890s	1910s	1930s	1940s	1950s	1960s	1970s	1980s	1990s
AP	0.60	1.00	0.15	0.16	0.12	0.10	0.13	0.07	0.07	0.06	0.04	0.03
NP	0.00	1.00	1.00	1.00	0.50	0.20	0.20	0.15	0.19	0.23	0.13	0.12
PP	0.00	0.00	1.00	0.40	0.67	1.00	2.00	0.29	0.27	0.09	0.04	0.03
NP+AP	0.00	0.00	0.75	0.80	0.60	0.50	0.75	0.73	0.67	0.60	0.50	0.35
NP+PP	1.00	1.00	1.00	0.24	0.19	0.25	0.23	0.22	0.36	0.28	0.09	0.11
NP+part	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00

We already know that some caution is warranted when comparing ratios for categories with large frequency discrepancies (cf. Chapter 4). If we look at the NP+AP category, for example, we see extremely high values for the type-token ratios, but the category never exceeded the threshold of 20 tokens overall. Still, it is interesting to see that the type frequency of the adjectival category has not followed the enormous increase in token frequency. In the most recent decennium, the 1,221 examples with AP intensifiers feature 39 different types (10 of which are hapaxes), whereas the NP category has 42 different types (16 of which are hapaxes) in 339 tokens. In Chapter 4, we saw that in present-day Netherlandic Dutch as well, the adjectival category was not characterised by a high degree of lexical variability, but by a number of highly frequent adjectival intensifiers. Many of these adjectival intensifiers were also found to be "all-round", flexible intensifiers, i.e. intensifiers that readily combine with a large variety of different verb types (cf. §5.3.2 infra). All of this seems to indicate that the frequency increase of the AP category is not chiefly driven by an expansion of the repertoire of adjectival intensifiers, but by an increase in the frequency of use of a (limited) number of highly frequent adjectival intensifiers. If we look at Table 5.5 again, the prepositional intensifiers have similar low TTR-scores, but only in the most recent decades. In fact, until quite recently, the prepositional category was not very well-represented at all. In the 1960s, there were only 15 tokens with a prepositional intensifier; this number jumped to 75 in the 1970s and by the 1990s there are 265. Given that the 265 tokens with PP intensifiers feature only 8 different types, a small set of intensifiers must be responsible for the increase. It is no coincidence that the take-off of the PP category more or less coincides with the introduction of the intensifying use of *te pletter* 'to smithereens'. The NP+PP category behaves more like the nominal intensifiers, and we see a remarkable increase in type frequency in the second half of the 20th Century. This category in particular has delivered a large proportion of hapaxes because both the NP-part and the PP-part of the intensifier

can display (small) variations, thus giving rise to new intensifiers (see also Ch3, §3.3.3 on the difference between surface form and lemma). The following subparagraphs provide a detailed description of the addition of new intensifiers to the range of attested slot fillers and their frequency developments, paying attention to both their formal and semantic properties. Based on our previous overall finding that the 1930s really seem to be a kind of breakthrough period for the construction, in terms of both token frequency and especially intensifier type frequency, we opt for a division between pre-1930s and post-1930s.

5.1.3.1 Early expansion: 19th Century until early 20th Century

Table 5.6 below lists the first attestations of all intensifiers per individual decennium and sorted by syntactic category.[56] The semantic categories that were established in Chapter 4 are provided between brackets (NEG: negatively connoted states, DIS: diseases, IP: inalienable possession, COL: colour terms, no label for the rest category of "other"). A full overview of the attested intensifiers per decennium can be found in Appendix V-2, and the English translations of all intensifiers are provided in the translation list at the beginning of this thesis.

Table 5.6. New additions to the intensifier repertoire per decennium (1830s to 1910s)

1830-1839:	AP: *dood* (NEG), *suf* (NEG), *ziek* (NEG) NP+PP: *een hart in het lijf* (IP)
1850-1859:	AP: *schor* (NEG) NP: *een bult* (DIS) NP+PP: *de ledematen uit de gewrichten* (IP), *de ogen uit het hoofd* (IP) NP+AP: *de longen te barsten* (IP)
1870-1879:	AP: *blind* (NEG), *halfdood* (NEG), *krom* (NEG) NP: *een ongeluk*, *een stuip* (DIS) PP: *ten doode* (NEG) NP+PP: *het vuur uit de sloffen* (IP) NP+AP: *de handen stuk* (IP), *de vingers krom* (IP), *de ziel dood* (IP)
1890-1899:	AP: *kapot* (NEG), *slap* (NEG), *halfgek* (NEG), *halflam* (NEG) PP: *in het zweet* (NEG) NP+PP: *het vuur uit de schoenen* (IP), *de zolen uit de sloffen* (IP) NP+AP: *de longen stuk* (IP), *de ogen blind* (IP), *de vingers moe* (IP)

[56] This does not mean that these are truly the very first occurrences in the intensifying fake reflexive resultative construction ever, just that these are the first attestations in *our corpus* (see Chapter 3 for the selected newspapers and decennia from the Delpher corpus).

1910-1919:	AP: *blauw* (COL), *bont en blauw* (COL), *groen en geel* (COL), *halfsuf* (NEG), *flauw* (NEG), *lam* (NEG) NP: *een aap*, *een beroerte* (DIS), *een koliek* (DIS), *stuipen* (DIS), *tranen* NP+PP: *een stuk in de kraag* (IP) NP+AP: *de vingers blauw* (IP), *de vingers groezelig* (IP), *het hoofd suf* (IP)

In the 1830s, we start out with some first examples of adjectival intensifiers from the semantic category of negatively connoted states (category I). While *ziek* 'sick' is no longer one of the top intensifiers in present-day Netherlandic Dutch (only 3 attestations in SoNaR-NL), *suf* 'drowsy' and *dood* 'dead' are the second and third most frequent intensifiers in the SoNaR corpus, respectively. As was mentioned in Chapter 3, §3.3.5.2, the early examples of *dood* 'dead' with the verb *werken* 'to work' are still somewhat reminiscent of the resultative construction, but we get some examples with experience verbs like *zich ergeren* 'to be annoyed' from the 1850s onwards.

(205) De koffijdrinkers zouden er **zich dood** over **ergeren**, wanneer de groenen en de zwarten gemeenschappelijk theewater hadden. (Delphcorp, 1850-1859)
the coffee-drinkers would it themselves dead over annoy [...]
'The coffee drinkers would be very annoyed when those who drank green tea and those who drank black tea got the same pot of tea water.'

It appears that Dutch also had a true translational equivalent of the English degree modifier *to death* in *ten doode*, see (206), but it was not frequent at all (only 5 tokens in total) and very limited in time: it appears to have been introduced in the 1870s but it was already gone by the 1930s. Its disappearance may have something to do with the archaic spelling and morphology: in present-day Dutch, *ten doode* would be *te dood/tot de dood*, which apparently did not make it into an intensifier.

(206) Deze Vorst zou zijn nieuwe Rijk kunnen regeeren met behulp van gepensionneerde Britsch-Indische ambtenaren, van welken er in Engeland legio **zich ten doode vervelen**. (Delphcorp, 1870-1879)
[...] many themselves to death bore
'This monarch could rule his new realm with the help of retired British-Indian officials, many of which are bored to death in England.'

Following several examples with *halfdood* 'half-dead', attested from the 1870s onwards, a lot of other negatively connoted states receive a variant with the modifier *half*, e.g. *halfgek* 'half-crazy', *halflam* 'half-lame', *halfsuf* 'half-drowsy', *halfslap* 'half-weak' and *halfziek* 'half-sick'.[57] It is worth mentioning that some of these adjectives, like *dood* 'dead' or *lam* 'lame',

[57] In some cases, the variant with *half* predates the use of the bare adjective as an intensifier in our corpus, e.g. *halflam* 'half-lame' was attested before *lam* 'lame'.

in their lexical sense do not really encode a "gradable" state. The fact that those adjectives can still be modified by *half* in this construction indicates that the original lexical semantics had already undergone considerable semantic bleaching. The intensifier *kapot* 'broken' was introduced in the 1890s, but it remains relatively infrequent until the mid-20th Century. Up until the 1950s, there are only 6 total attestations with *kapot* 'broken', which stands in stark contrast to it currently being the fourth most common intensifier in the SoNaR-NL corpus. There are several other intensifiers from the first semantic category that are not yet attested in the construction in the 19th or early 20th Century, e.g. *uit de naad* 'out of the seam', *wild* 'wild', *wezenloos* 'blank/vacant', *te pletter* 'to smithereens', etc. Perhaps most striking is the fact that there is no sign yet of *rot* 'rotten', which is the most frequent intensifier in present-day Dutch in the SoNaR-NL corpus.

Throughout the 19th and early 20th Centuries, we occasionally find NP+PP intensifiers belonging to the semantic category of inalienable possession (category III). These intensifiers appear to have been introduced as part of a specific verb-intensifier combination rather than as "productive" intensifiers, because it is the entire collocation that is adopted by other speakers. That is the case for, for example, *zich de ogen uit het hoofd schamen* 'to embarrass oneself the eyes out of the head' (first attestation in the 1850s), and *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' (first attestation in the 1870s). Curiously, in the first attestation of *de ogen uit het hoofd* 'the eyes out of the head', the NP and PP are separated by the verb, which may suggest that both elements are still processed individually and not experienced as a fixed chunk at the time. Also note the play on words in (208), in which the element of slippers in the intensifier is taken up again later in the sentence as well.

(207) Waarlijk, zij mag **zich** wel **de oogen schamen uit het hoofd**. De Nieuwsbode, haar orgaan, verhoogt haar aanzien niet. (Delphcorp, 1850-1859)
truly, she may herself really the eyes embarrass out of the head [...]
'Truly, she should be very ashamed. De Nieuwsbode, her mouthpiece, does not add to her prestige.'

(208) En of we met Juni **ons** nu **het vuur uit de sloffen loopen**, of "op onze slofjes" thuis blijven, voor de uitkomst maakt dat toch geen noemenswaardig verschil. (Delphcorp, 1870-1879)
and if we in june ourselves now the fire out of the slippers run [...]
'Whether we run our socks off or stay at home "in our slippers" in June, it won't make much of a difference in the end.'

Over time we come across a couple of isolated examples in which these intensifiers were extended to other verbs, as in (209), but the original verb-intensifier combination remains the most frequent even in the most recent decennia (see §5.3.2 on the relaxation of collocational constraints and collocational expansion).

(209) De menschen werden, nadat ze **zich** eerst **de oogen uit hun hoofd** hadden **verbaasd**, wanhopig. (Delphcorp, 1910-1919)
[...] after they themselves first the eyes out of the head had amazed [...]
'The people, after they got over the first shock, became desperate.'

This suggests that the expression is perhaps not so much "fixed" in the strict sense of the word, but it definitely has a certain conventional status. This is in line with Zeschel's (2012: 25) definition of fixed expressions as "pre-assembled holistic units that are not assembled from scratch [but] not necessarily frozen in the sense of 'not tolerating any lexical subtitution' whatsoever". This "lexical substitution" does not only cover the occasional extension of these intensifiers to new verbs, but also small variations in the intensifier itself. While the use of *het vuur uit de sloffen* 'the fire out of the slippers', for example, was restricted to the verb *lopen* 'to run' until the 1950s (cf. infra), we already see some small variation soon after it was introduced, e.g. *schoenen* 'shoes' instead of *sloffen* 'slippers' in (210).

(210) Daar is ten eerste de grootmoeder, ging hij voort, die **zich het vuur uit de schoenen loopt**, om den kleinen jongen zoo dikwijls mogelijk te zien. (Delphcorp, 1890-1899)
[...] who runs herself the fire out of the shoes [...]
'First of all we have the grandmother, he proceeded, who runs around like crazy to see the little boy as often as possible.'

The other NP+PP or removal intensifiers that have some frequency in present-day Dutch, viz. *de longen uit het lijf* 'the lungs out of the body' or *de benen uit het lijf* 'the legs out of the body', were not yet attested before the 1940s.

Compared to the one example of an NP+AP intensifier (i.e. *de nieren los* 'the kidneys loose') in the SoNaR-NL corpus, we find some more variation in this category in Delphcorp. Like most NP+PP intensifiers, the NP generally denotes an inalienable body-part and the AP adds that the body-part is affected in one way or another (category III). The NP+AP intensifiers are generally context-specific (and infrequent) and there is a clear conceptual link between the body-part in the intensifier and the verb it occurs with, see (211) and (212) (see also §5.3.2 below).

(211) De dochter, Mariette heet ze, **schreit zich de oogen blind**, zoo hing ze aan de moeder! (Delphcorp, 1890-1899)
the daughter, mariette is named she, cries herself the eyes blind [...]
'The daughter, Mariette, cries her eyes out, that's how clingy she is with her mother.'

(212) De bewondering van alle lieden van kennis en opvoeding, die kregen zij niet. Daarentegen zou de domme schare **zich de handen stuk klappen**. En dat is ook iets. (Delphcorp, 1870-1879)
[...] would the dumb crowd itself the hands broken clap [...]
'They did not receive any admiration from the well-educated, but the rabble would clap with fervour. It's something.'

In this period we also find some first examples with colour adjectives that would grow out to be very frequent as part of conventional collocations in present-day Dutch, e.g. *zich blauw betalen/ergeren* 'to pay/annoy oneself blue' and *zich groen en geel ergeren* 'to annoy oneself green and yellow'. Both colour intensifiers were introduced in the 1910s as hapaxes, but, curiously enough, not with the verbs they would eventually form a strong collocation with. While *groen en geel* 'green and yellow' will be exclusively paired up with *zich ergeren* 'to be annoyed' from the 1930s onwards, its first occurrence in the 1910s is actually with the verb *zich vervelen* 'to be bored'. *Blauw* 'blue' does display some use outside of its collocations in present-day Dutch, but it is still somewhat surprising that the first attestation is with *wachten* 'to wait', a rather infrequent verb in the construction overall. Given that we are now so familiar with the conventional collocations with *groen en geel* 'green and yellow' and *blauw* 'blue', both examples below sound rather odd from a present-day perspective.

(213) Jullie moogt op je veranda's zitten en **je groen en geel vervelen** en den saaien Zondag verwenschen. (Delphcorp, 1910-1919).
[...] and bore yourselves green and yellow [...]
'You can sit on your porches, bored out of your mind, while cursing the boring Sunday.'

(214) Is er dan in die vijf maanden niets aan gedaan en moeten de werklieden **zich** dan nu maar weer **blauw wachten**? (Delphcorp, 1910-1919)
[...] the workmen themselves then now but again blue wait
'So in those five months nothing was done about it, and the workmen should just keep waiting?'

The category of colour terms was not very well-represented in our present-day SoNaR-NL corpus, but we will see in the next section that it is an important source for analogical extensions in the later decades in the 20th century.

The semantic category of diseases is still relatively underrepresented in this first period under investigation as well. In the early 20th Century, the intensifiers *een koliek* 'a colic', *een stuip* 'a spasm' and *een beroerte* 'a stroke' are sporadically attested, and we also find some examples with *een bult* 'a hump', all with the verb *lachen* 'to laugh'.

(215) Vol belangstelling werd hij aangehoord, totdat Smit hem vroeg: Heb je **je** niet **een beroerte gelachen**? (Delphcorp, 1910-1919)
[...] have you yourself not a stroke laughed
'People listened to him attentively until Smit asked him: Did you not find it hilarious?'

(216) Er zijn menselijke schepsels [...] die met strakke gezichten kunnen luisteren naar een grap, waarover een ander **zich een stuip** zou **lachen**. (Delphcorp, 1910-1919)
[...] about which another himself a spasm would laugh
'There are human creatures who maintain a straight face when listening to a joke that would crack other people up.'

The "taboo" disease terms which have a higher expressive force (i.e. often informal terms for diseases that are, or used to be, lethal) like *de pokken* 'the small-pox', *de pleuris* 'the pleurisy', *de tering* 'the consumption', etc. were not used in the construction yet, nor were any of the fictitious diseases. Given the large variety of such diseases in present-day Dutch, the type explosion in the category of diseases must be of a more recent date (cf. infra).

In the remainder category, there are also a number of intensifiers that are relatively frequent in present-day Dutch but still absent at this point in time, e.g. *een hoedje* 'a little hat' and *een slag in de rondte* 'a punch around'. *Een aap* 'a monkey' is attested in the corpus from the 1910s onwards. This is interesting given that it was completely absent in the present-day Dutch SoNaR-NL corpus, while being relatively frequent in Belgian Dutch (cf. infra).

(217) Daar begint de match, en het was waarlijk, om **je een aap** te **lachen**. (Delphcorp, 1910-1919)
[...] to yourself a monkey to laugh
'So the game began and it was truly hilarious.'

In sum, the intensifier slot in the construction already underwent quite an expansion between 1830 and 1919, gradually attracting new intensifiers from different syntactic and semantic categories. While it started out with a number of negatively connoted adjectives, which already appear to have developed an intensifying function by the mid-19th Century, we soon saw the addition of some first NP intensifiers like *een ongeluk* 'an accident' and NP+PP intensifiers (or verb-intensifier combinations) like *(zich) het vuur uit de sloffen (lopen)* '(to run oneself) the fire out of the slippers', some of which will become highly frequent in present-day Dutch. Others, like *een bult* 'a hump', *een kriek* 'a hump' or *een stuip* 'a spasm' pop up in the construction occasionally throughout time and can still be used in present-day Dutch, but they have never exceeded the threshold of five occurrences. Additionally, we found one example of an intensifier that had a short lifespan, i.e. *ten doode* 'to death' between the 1870s and 1919. However, there is still a big difference between the situation around the late 1910s and the present-day situation that was described in Chapter 4. Several of the intensifiers in the semantic category of negatively connoted states that are moderately to very frequent in present-day Dutch had not been introduced yet at the time, and the type explosion in the category of disease terms also appears to be of a more recent date. We will zoom in on this more recent period in the next subparagraph.

5.1.3.2 Recent expansion: after the 1930s

While the construction had already been gradually increasing its token and type frequencies in the previous period under investigation, both the token and type frequency curves in Figure 5.1 and Figure 5.3 started increasing at a steeper rate around

the mid-20th Century. The parallelism between the token and type frequency curves suggests that the general increase in the frequency of use of the construction is concomitant with an increasing expansion of the range of different verbs and intensifiers. In §5.1.1, we found that the token frequency increase is only partly explained by the construction being extended to *new* verb types; the general expansion is primarily carried by a number of highly frequent verbs. Based on the information on token and type frequency in Figure 5.12, we have proposed that this may also be the case for the intensifier slot. Figure 5.13 captures the development of the twenty most frequently used intensifiers in the entire data set for the period 1930-1995 (based on the summed frequencies over all decennia).

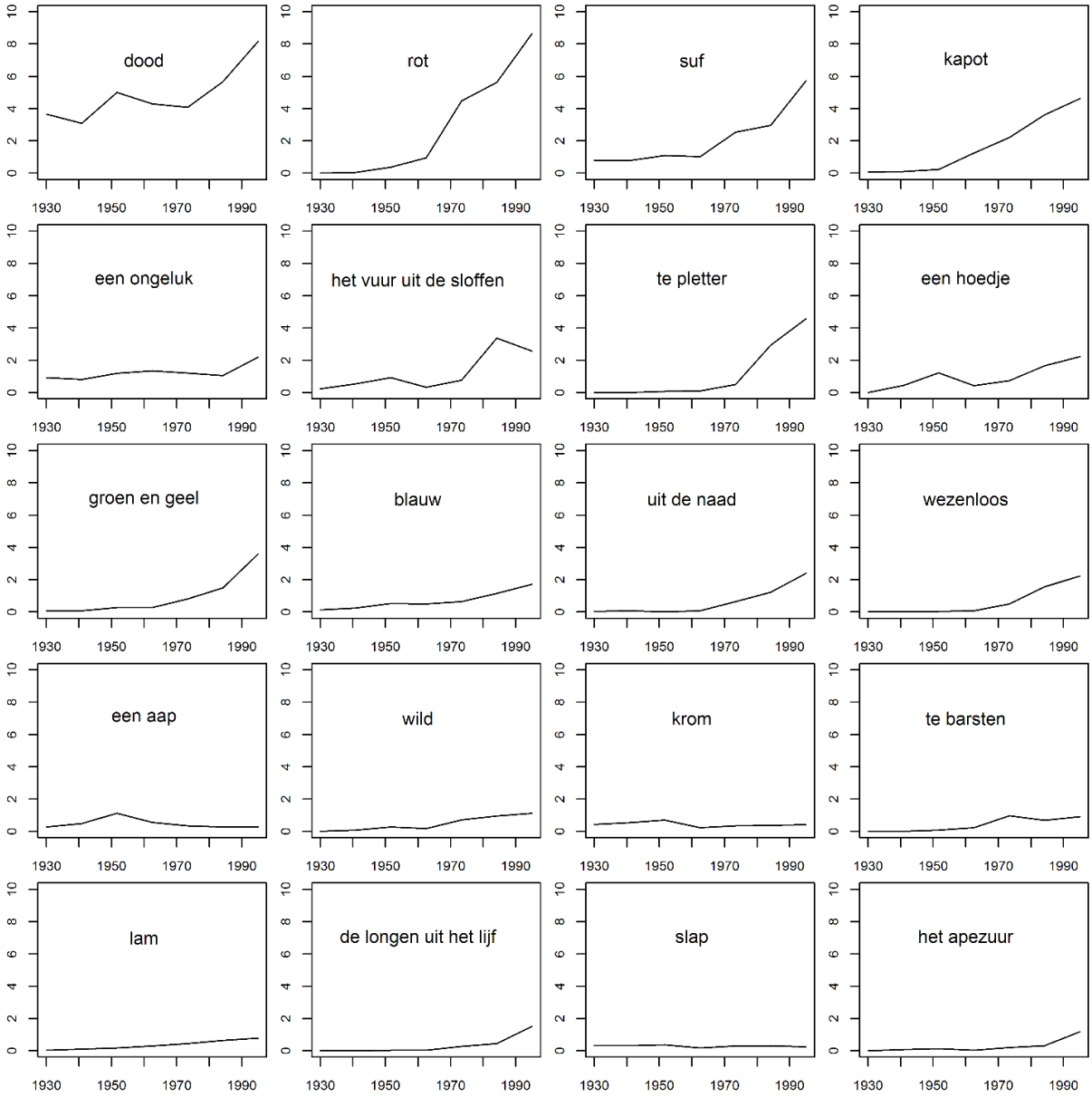


Figure 5.13. Normalised token frequency development of the 20 most frequent intensifiers (1930s-1990s)

Indeed, the figure shows that the overall frequency increase of the construction was not only carried by a number of highly frequent verbs, but also by a number of intensifiers. Considering the steep increases in the top panels, it appears that the second half of the 20th Century was a crucial breakthrough point for most of the top ten to fifteen intensifiers as we know them today (see Chapter 4). Given that a lot of these top intensifiers are adjectives, e.g. *rot* 'rotten', *suf* 'drowsy', *kapot* 'broken', *groen en geel* 'green and yellow' and *wezenloos* 'blank/vacant', this confirms our earlier claim that the enormous token frequency expansion that was attested for the adjectival category in particular was spearheaded by a limited number of highly frequent adjectival intensifiers. The graph also shows that there is considerable variation between the development of the individual intensifiers. While most intensifiers display a noticeable upward trend, others appear to be on a declining trajectory. The intensifier *het vuur uit de sloffen* 'the fire out of the slippers' was increasing up until the 1980s, but shows a sudden dip in the 1990s. On the basis of the current data, we cannot say whether this announces a lasting decreasing trend or whether it is but a minor temporary setback, similar to the ones that are observed for *een hoedje* 'a little hat' around the 1960s or *dood* 'dead' around the 1970s. A better example of an intensifier that has outlived its fashion is *een aap* 'a monkey': while it was still used in 34 tokens in the 1950s, it dropped down to 17 in the 1960s, 10 in the 1970s and 8 in the 1980s and 1990s, eventually leading to zero in the SoNaR-NL corpus. If *een aap* has disappeared as an intensifier in Netherlandic Dutch, while still being used in Belgian Dutch, this adds some strength to the hypothesis that Belgian Dutch is – in some respects – still somewhat more "archaic" than Netherlandic Dutch (cf. §2.1.1.1). Although it is not part of the top twenty intensifiers and hence not included in Figure 5.13, we see a similar evolution for *ziek* 'sick'. Being one of the earliest attested intensifiers in the construction, it peaked in the 1950s with 18 tokens but has been decreasing in frequency ever since. Again, its use as an intensifier is preserved better in Belgian Dutch than in Netherlandic Dutch (cf. Chapter 4). Finally, there are some intensifiers that show a certain stability in their frequency development or have (slightly) increased only very recently, such as *een ongeluk* 'an accident', *lam* 'lame' and *slap* 'weak'.

Again, we follow Hilpert (2015a) in representing the share of the top twenty intensifiers set off against the remaining 190 intensifiers over time, see Figure 5.14. The upper panel shows that there is an increase of both frequent and infrequent intensifier types over time, but the lower panel of the graph suggests that the share of high-frequency intensifier types becomes more substantial over time. One possible reason behind this development is that many of the intensifiers that are in the top twenty are recent success stories. For example, *rot* 'rotten', *kapot* 'broken', *te pletter* 'to smithereens', *uit de naad* 'out of the seam' or *een hoedje* 'a little hat' barely had any attestations before the 1950s, so their impact on the total data set was very limited before then. While this is indeed an interesting finding that will be shown to have an important impact on the network representation (which will be discussed in §5.4), there are a lot of interesting

developments in the overall use and semantic range of the intensifying fake reflexive resultative construction that cannot be directly gleaned from this graph.

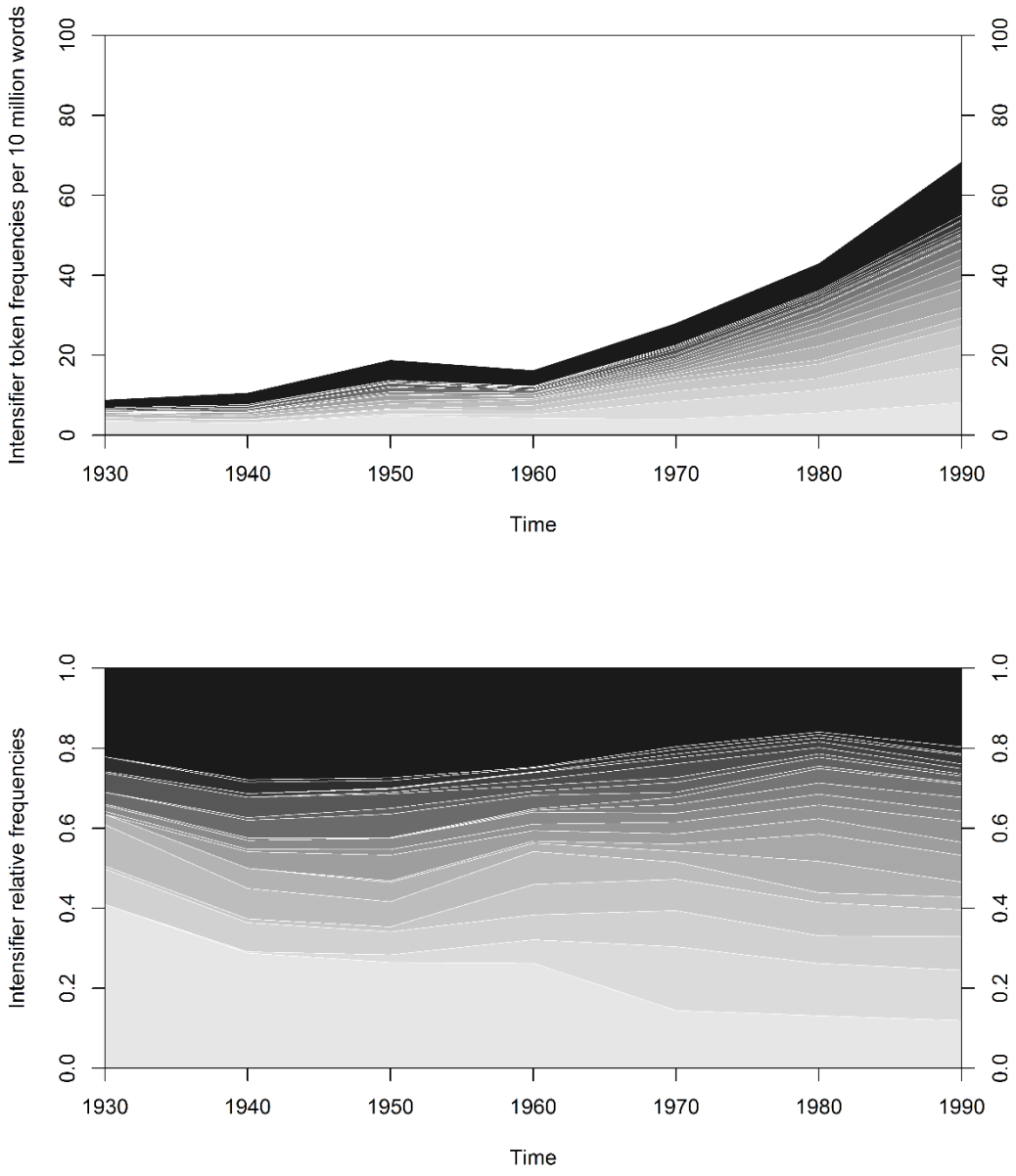


Figure 5.14. Frequency developments of 20 frequent intensifier types versus 190 infrequent intensifier types. (upper panel: normalised frequency development, lower panel: relative frequency development)

Following Table 5.6, we continue to list the new additions to the intensifier slot for the more recent decennia in Table 5.7.

Table 5.7. New additions to the intensifier repertoire per decennium (1930s to 1990s)

1930-1939:	AP: *gek* (NEG), *halfslap* (NEG), *halfziek* (NEG), *kleurenblind* (NEG), *ziek en weer gezond* (NEG), *zwart* (COL) NP: *een hartverlamming* (DIS), *een hoed*, *een kriek* (DIS), *een mik*, *een pukkel* (DIS) PP: *uit de naad* (NEG) NP+PP: *de benen uit het lid* (IP) NP+AP: *de buik rond* (IP)
1940-1949:	AP: *bewusteloos* (NEG), *geel en groen* (COL), *groen* (COL), *groen en blauw* (COL), *groen en grijs* (COL), *leeg* (NEG), *ongelukkig* (NEG), *rood, wit en blauw* (COL), *rot* (NEG), *stuk* (NEG), *wild* (NEG) NP: *de blaren* (DIS), *de pest* (DIS), *een aapje*, *een hoedje*, *een rotje*, *het apezuur* (DIS), *tranen met tuiten* PP: *over de kop* (NEG), *uit het lid* (IP) NP+PP: *de nekspieren uit het lid* (IP), *het vuur uit de schenen* (IP), *het vuur uit de voetbalschoenen* (IP) NP+AP: *de handen kapot* (IP), *de longen leeg* (IP) NP+part: *de ogen uit* (IP)
1950-1959:	AP: *grasgroen* (COL), *lens* (NEG), *rond* (NEG), *wezenloos* (NEG), *zenuwziek* (NEG) NP: *de mazelen* (DIS), *de pokken* (DIS), *de stuipen* (DIS), *de tranen*, *een aanp*, *een breuk* (DIS), *een halve beroerte* (DIS), *een kokosnoot*, *een punthoofd* (DIS?), *een zuurstok*, *het hoedje*, *het rambam*, *het zuur* (DIS) PP: *te barsten* (NEG), *te pletter* (NEG) NP+PP: *de benen PREP het lijf* (IP), *de blaren in de handen* (IP), *de longen uit het lijf* (IP), *het schuim op de mond* (IP), *het vuur uit de pen*, *het vuur uit de vingers* (IP), *het vuur uit de rennerssloffen* (IP) NP+AP: *de handen blauw* (IP), *de hersenen suf* (IP), *de ogen zat* (IP)
1960:1969:	AP: *beroerd* (NEG), *grijs* (COL), *ongans* (NEG/DIS), *paars* (COL), *schel* (NEG), *stom* (NEG) NP: *de blubber, de krampen* (DIS), *de rambam*, *de zenuwen* (DIS), *het apelazerus* (DIS), *het lazerus* (DIS) NP+PP: *de adem uit de longen* (IP), *de poten PREP het lijf* (IP), *het vuur uit de schaatsen* (IP), *het vuur uit de slofjes* (IP) NP+AP: *de hakken scheef* (IP), *de nagels blauw* (IP), *de vingers wond en rond* (IP)
1970-1979:	AP: *bleek* (NEG), *bloot* (NEG), *geel* (COL), *halfkapot* (NEG), *klem* (NEG), *lazerus* (NEG/DIS), *scheef* (NEG), *tureluurs* (NEG), *verrot* (NEG) NP: *de klere* (DIS), *de kolere* (DIS), *de/het pleuris* (DIS), *de stuipjes* (DIS), *een barst*, *een loei*, *een puist* (DIS), *een rotberoerte* (DIS), *een slag in de rondte*, *een slaghoedje*, *het laplazerus* (DIS), *het schompes* (DIS), *krampen* (DIS) PP: *in de poeier* (NEG), *uit het lood* (NEG) NP+PP: *de benen PREP het gat* (IP), *de blaren op de tong* (IP), *de naad uit het lijf* (IP), *de ogen uit de kassen* (IP), *het licht uit de ogen* (IP), *het schuim op de mond* (IP), *het schuim op de ziel* (IP), *het schuim op de hiel* (IP), *het vuur uit de spikes* (IP), *het vuur uit de sportsloffen* (IP), *de zolen PREP de voeten* (IP) NP+AP: *de poten kapot* (IP)

1980-1989:	AP: *azuur-blauw* (COL), *grijs en groen* (COL) NP: *de pestpokken* (DIS), *een deuk*, *het leplazerus* (DIS), *het ongans* (DIS) PP: *te blubber* (NEG), *uit de naden* (NEG) NP+PP: *de benen uit de naad* (IP), *de blaren op de hakken* (IP), *de blaren PREP de voeten* (IP), *een stuk in de hakken* (IP), *het vuur uit de spaken* (IP) NP+AP: *de oren rood* (IP)
1990-1995	AP: *apelazerus* (NEG/DIS), *blauw en groen* (COL), *gaar* (NEG), *laveloos* (NEG), *paars en groen* (COL), *rood en groen* (COL) NP: *n° slagen in de rondte*, *blaren* (DIS), *de griebels* (DIS?), *de hik* (DIS), *de takken* (DIS), *de tering* (DIS), *de tering-takke* (DIS), *de vinketering* (DIS), *een blauw hart* (DIS?) PP: *in pust* (NEG/DIS) NP+PP: *blaren op de tond* (IP), *de bril van het hoofd* (IP), *de longen uit de balg* (IP), *de voeten PREP het lijf* (IP), *een stuk in de kont* (IP), *het snot voor de ogen* (IP), *het vuur uit de molières* (IP) NP+AP: *de heupen stuk* (IP), *de vingers beurs* (IP), *het hoofd gek* (IP), *ledematen blauw* (IP)

If we compare the list above to Table 5.6, it is obvious that the lists start getting longer, as is to be expected based on the type frequency development in Figure 5.12. Some of the intensifiers are here to stay and will spread through the linguistic community, but many of these are creative one-offs and are not likely to gain much ground. There may be a correlation between the frequency increase of some highly frequent intensifiers and the sudden boost of new types and hapaxes in the second half of the 20th Century. As a number of intensifiers are rapidly gaining in frequency, they may also be losing their expressive force, which in turn may trigger an influx of new, creative formations (see Ch2, §2.3 and §5.5.1 infra). In Chapter 4, we saw that the majority of the infrequent or hapax intensifiers in present-day Dutch belong to one of the pre-established semantic categories. In Table 5.7 as well, we find that most creative coinages fit into one of those semantic categories. This is in accordance with the hypothesis that new types generally reflect the semantic space that is covered by the already established types (Zeschel 2012: 185). Some categories appear to be more prolific in spawning new intensifiers than others. In particular, the group of intensifiers denoting diseases, which was not very populated yet in the early 20th Century, has undergone a considerable expansion from the 1940s onwards. In the 1940s and 1950s, we get some examples of the taboo diseases like *de pest* 'the plague' and *de pokken* 'the smallpox', as well as a first example of a fictitious disease, *het apezuur* 'lit. the monkey-heartburn'.

(218) En wat wou u? 42 uur werken? Nee, 4114 uur! Ja, wij zullen **ons de pokken werken**! (Delphcorp, 1950-1959)
[...] yes, we will ourselves the smallpox work
'So what did you want? 42 hours of work? No, 4,114 hours! Sure, we'll work our butts off!'

From the mid-20th Century, we observe an explosion of both real and fictitious diseases, e.g. *het lazerus* 'the leprosy', *de/het pleuris* 'the pleurisy', *de klere/kolere* 'the cholera', *de tering* 'the consumption', *de pestpokken* 'the plague-smallpox', *het apelazerus* 'fictitious disease', *het laplazerus* 'fictitious disease', *het schompes* 'fictitious disease', *de tering-takke* 'fictitious disease', *de vinketering* 'fictitious disease', etc. By the 1990s, we have an inventory of approximately 20 different diseases that are used as intensifiers in the construction.

(219) Heb ik me hier voor 20 jaar **het schompus gewerkt** om op zo'n rotfestivalletje te spelen? (Delphcorp, 1970-1979)
have I myself here for 20 years the schompus worked [...]
'Have I been working my butt off for 20 years to play on such a miserable little festival?'

(220) Die hebben zich allemaal **de tering-takke geërgerd** aan het feit dat die kutmoffen wonnen. (Delphcorp, 1990-1995)
they have themselves all the consumption-piles annoyed [...]
'They were all so annoyed by the fact that the fucking Germans won.'

Of course, it is not implausible that these intensifiers were already used in the construction before the mid-20th Century, but that they were at the time still considered inappropriate in newspaper language (see §5.5.2 for some discussion).[58] This influx of disease intensifiers is probably what boosted the lexical variety of the category of nominal intensifiers (cf. supra). At the same time, we find some conversions of the NP diseases to other syntactic categories (see also Ch3, §3.3.3). While *het apelazerus* was introduced in the 1960s, we find an AP variant *apelazerus* in the 1990s. Similarly, *de pleuris* 'the pleurisy' is first attested in the 1980s but we find AP and PP variants *pleuris* and *te pleuris* in the present-day SoNaR data (cf. Chapter 4).

The category of colour terms as well starts to expand its range around the mid-20th Century. From the 1940s onwards, we already find different colour combinations, e.g. *groen en grijs* 'green and grey', *groen en blauw* 'green and blue', *geel en groen* 'yellow and green', *grijs* 'grey', *paars* 'purple' and even *zwart* 'black'. Despite the lexical variety, most colour intensifiers appear to be limited to the verb *zich ergeren* 'to be annoyed' (cf. §5.3.2 and §5.4 infra).

[58] It is further striking to note that these disease intensifiers mainly denote diseases that are effectively eradicated in the western world, and, in that light, may have shed some of their "taboo" connotation, e.g. the plague, tuberculosis, cholera, etc. Certain (deadly) diseases that are still very present today, like cancer or aids, do occur in the construction in informal language use but they are not (yet) attested in newspapers. *Jezus christus de hond piepte ineens heel hard en ik* ***schrok me de kanker*** *en nu ben ik mijn telefoon kwijt* (Twitter 15/04/2013) 'Jesus Christ, the dog suddenly squeeked really loud and I was so startled (lit. startled myself the cancer) and now I've lost my phone'; *Vandaag kzat rustig in me kamer te zitten opeens me pa zet harde surinaamse muziek aan* ***schrok me de aids.*** (Twitter 06/07/2012) 'Today I was chilling in my room and suddenly my dad turns on this loud Surinamese music, I was so startled (lit. startled myself the aids)'

(221) Er zijn 1200 van deze lieden in den lande, die er **zich groen en grijs** aan **ergeren**, dat zij geen toewijzing van de distributie ontvangen. (Delphcorp, 1940-1949)
[...] who it themselves green and grey about annoy [...]
'There are 1,200 people in this country who are very annoyed that they do not receive any allocation from the distribution centre.'

(222) Al jarenlang **erger ik mij paars** aan het werkelijk waardeloze filmaanbod in deze stad. (Delphcorp, 1970-1979)
for years-long annoy I myself purple [...]
'For years, I've been very annoyed by the truly worthless film offer in this city.'

The category of negatively connoted states was already well-represented in terms of intensifier types in the older period discussed in the previous section. The 19th and early 20th Century saw the introduction of a number of intensifiers that would develop into extremely frequent, all-round intensifiers in present-day Dutch. As we will see below, these intensifiers continued to increase in frequency in the second half of the 20th Century. We also noted that a number of intensifiers which are frequently used in present-day Dutch, were still unattested before the 1930s. This is the case for *rot* 'rotten' (the most frequent intensifier in present-day Netherlandic Dutch), which was introduced as a hapax in the 1940s.

(223) God, zei Fred, het is om **je rot** te **lachen**. Laten we nog een borrel nemen. (Delphcorp, 1940-1949)
god, said fred, it is to yourself rotten to laugh [...]
'God, Fred said, isn't it hilarious? Let's have another drink.'

Others include *wild* 'wild' (1940s), *te pletter* 'to smithereens' (1950s), *wezenloos* 'blank/vacant' (1950s) and *een slag in de rondte* 'a punch around' (1970s).

(224) Die vent **schrikt zich wild** als Nico hem aanhoudt. (Delphcorp, 1940-1949)
that man startles himself wild [...]
'The man is very startled when Nico stops/arrests him.'

(225) Het geschenk, dat hij haar had willen aanbieden en waarvoor hij **zich** dagenlang **wezenloos** had **getrapt** [...]. (Delphcorp, 1950-1959)
[...] he himself days-long vacant had pedaled [...]
'The gift that he had wanted to offer her, and for which he had cycled his butt off for days...'

With respect to the third category of intensifiers, i.e. inalienable possession, we observe that *het vuur uit de sloffen* 'the fire out of the slippers' gains in frequency, while extending its use to other verbs besides *lopen* (see (226) and (227)) and triggering a number of analogical extensions like het *vuur uit de schenen* 'the fire out of the shins', *het vuur uit de voetbalschoenen* 'the fire out of the soccer shoes', *het vuur uit de rennerssloffen* 'the fire out of the cycling slippers', *het vuur uit de spaken* 'the fire out of the spokes'. It is possible that the original form *het vuur uit de sloffen* had lost some of its expressive force, which

encouraged language users (in this case journalists) to create a variant that is more likely to be noticed by the reader, while also providing a better fit in specific contexts.

(226) De "Vrienden van het Frysk Orkest" **werken zich het vuur uit de sloffen** om het voorlopig benodigde bedrag bij elkaar te krijgen. (Delphcorp, 1950-1959)
the friends of the frisian orchestra work themselves the fire out of the slippers [...]
'The "Friends of the Frysk Orkest" work hard to raise the currently necessary amount.'

(227) Zelfs déze activiteit schijnt toch passief, vergeleken bij de twee en twintig voetballers plus één scheidsrechter, die **zich het vuur uit de voetbalschoenen draven en schoppen**. (Delphcorp, 1940-1949)
[...] who themselves the fire out of the soccer-shoes trot and kick
'Even this activity appears to be passive, compared to the twenty-two soccer players and one referee that run and kick their hearts out.'

Like in the 19th and early 20th Century, the new additions to the category of intensifiers involving inalienable possession are mostly context-specific and infrequent. Two exceptions are the NP+PP intensifiers *de longen uit het lijf* 'the lungs out of the body' and *de benen uit het lijf* 'the legs out of the body', both introduced in the 1950s. While *de benen uit het lijf* 'the legs out of the body' mostly gains ground as part of the conventional collocation *zich de benen uit het lijf lopen* 'to run oneself the legs out of the body', *de longen uit het lijf* 'the lungs out of the body' enjoys some combinatorial flexibility (cf. infra).

(228) We **liepen ons de benen uit het lijf** om zo snel mogelijk uit die ijlheid te raken. (Delphcorp, 1950-1959).
we ran ourselves the legs out of the body [...]
'We hurried to get away from the thin air as quickly as possible.'

(229) Ook kan niemand u horen, al **schreeuwde** u **zich de longen uit het lijf**. (Delphcorp, 1950-1959)
[...] yelled you yourself the lungs out of the body
'Nobody can hear you, even if you shouted from the top of your lungs.'

Finally, the 1940s also introduce the collocation *zich een hoedje schrikken* 'to startle oneself a little hat', which will quickly become one of the strongest conventional collocations in the construction. Curiously, before the first attestation of *een hoedje* 'a little hat', the data already contained one example with the intensifier *een hoed* 'a hat' in the 1930s, also with the verb *schrikken* 'to be startled'.

(230) Wat is er? vroeg ik, en ik **schrok me een hoed**. (Delphcorp, 1930-1939)
[...] and I startled myself a hat
'What is wrong? I asked, and I was very startled.'

(231) Hij weet zowaar niet, wat hem overkomt en **schrikt zich 'n hoedje**. (Delphcorp, 1940-1949)
[...] and startles himself a little hat
'He has no idea what is happening to him and it scares him a lot.'

While all of this undeniably points to a continued expansion of the intensifier repertoire in the second half of the 20th Century, it does not yet explain the rise-and-fall pattern in type frequency that was attested in Figure 5.12. Now that we have an overview of the new intensifiers that were added in each of the decennia, we may find an explanation for the attested fluctuation. We hypothesised earlier that the 1950s and 1970s, for whatever reason, may have been characterised by an exceptional boost in creativity with respect to the creation of new intensifiers. If we look at Table 5.7 above, we find that the lists of new additions for the 1950s and the 1970s do appear to be longer than for the surrounding decades. Given that creative instances are often operationalised as hapax legomena in corpus investigations (Zeschel 2012: 185, 228), this is where we need to turn our attention to. Table 5.8 shows the impact of hapaxes on the type frequency development by providing the ratio of hapaxes to total types, as well as the type frequencies if the hapaxes are taken out of the equation.

Table 5.8. The hapax/type ratio and type frequency for non-hapax intensifiers (1930s to 1990s)

	1930s	1940s	1950s	1960s	1970s	1980s	1990s
Hapax/type ratio	0.51	0.39	0.49	0.38	0.44	0.28	0.34
Type frequency (n>1)	18	35	36	40	54	55	76
Real* hapax count	7	5	16	6	18	6	22

The 1950s and the 1970s appear to have a higher proportion of hapaxes than their surrounding decades – together, they account for 34 out of the 89 hapaxes in the entire data set – which corroborates our hypothesis that language users were exceptionally inventive in these two decades. If we do not take the hapax occurrences into consideration, we do see a steady increase in type frequency from the 1930s to the 1990s. Of course, not all hapaxes are necessarily new, creative coinages. There are several hapax intensifiers that could be described as "occasional visitors" in the construction: they show up in the intensifier slot every now and then, without ever gaining any consistent frequency, e.g. *een stuip* 'a spasm', *een koliek* 'a colic', *schor* 'hoarse', etc. If we ignore these instances for the time being and only count the real hapaxes, we still find a lot of truly unique intensifiers in the 1950s and the 1970s. Some examples are given below in (232) to (235).[59]

[59] By real hapax count, we mean that we are only counting the hapax legomena that are also hapaxes in the entire data set, not just hapaxes in this particular decennium.

(232) Minister Moedweg **schrok zich een kokosnoot** en gooide zijn waterkaraf om. (Delphcorp, 1950-1959)
minister moedweg startled himself a coconut [...]
'Minister Moedweg was very startled and knocked over his water carafe.'

(233) Die mijnheer **schrok zich** natuurlijk **een zuurstok** toen hij dit las. (Delphcorp, 1950-1959)
that mister startled himself of course a stick of rock when he this read
'That sir was obviously very startled when he read this.'

(234) Oj oj oj, een film om **je bloot** te **lachen**, goed gek en lekker pikant! (Delphcorp, 1970-1979)
[...] to yourself naked to laugh [...]
'Oh oh oh, a truly hilarious movie, totally bonkers and smoking hot!'

(235) De anderen **lachten zich een barst** als hij vertelde hoe de inspecteur van Staatstoezicht telkens opnieuw voor de gek gehouden werd. (Delphcorp, 1970-1979)
the others laughed themselves a crack [...]
'The others had a good laugh when he told them how the Inspector of State was made a fool of time after time.'

At the same time, the 1950s and the 1970s also marked the introduction or the take-off of some intensifiers that would stick around and become one of the top twenty intensifiers in present-day Dutch (e.g. *rot* 'rotten', *te pletter* 'to smithereens', *kapot* 'broken', etc.).

This section already briefly touched upon the collocational behaviour of some specific slot fillers when we analysed the contexts in which new verbs and/or intensifiers emerged. In the next section, these collocational patterns will be analysed in more detail.

5.2 Collocational patterns: expansion and conventionalisation

In the previous section, it was established that both the verb slot and the intensifier slot have been gradually increasing their respective ranges of attested slot fillers since the early 19th Century, with a clear boost in the second half of the 20th Century. We also found that, once the construction started to gain some frequency, both slots started to display the typical Zipfian distribution (Ellis & Ferreira-Junior 2009, Gries 2012), in which a limited set of highly frequent items already account for a large part of the data. Interestingly, although the construction has extended its use to new verb types, there was remarkable consistency in the verbs that were featured in the top ten of most frequently used verbs. In the intensifier slot we also find a couple of long-time developers in the top ten, but there has been noticeably more variation in the most frequently used intensifiers over time. This indicates that whereas there is apparently little change in the verbs that are particularly suited for intensification, there are some important shifts in the lexical items that can serve as potential intensifiers in the construction. However, we

may wonder to what extent we can really tease the development of the two central slots of the construction apart. If we look at the parallelism of the curves in Figure 5.3 and Figure 5.6, both slots appear to have developed in a largely parallel fashion. Moreover, Chapter 4 showed that there are important interactions between the two open slots of the construction in present-day Dutch and that specific verbs and intensifiers often enter into conventional combinations. If certain verbs and intensifiers have both been part of the same conventional (or conventionalising) combination, their frequency development is evidently tightly interrelated.

One way of tracking the emergence of such conventional collocations is through the diachronic application of the covarying collexeme analysis (see, e.g., Stefanowitsch & Gries 2003, 2005, Gries & Stefanowitsch 2004a, and Chapter 4 for a detailed explanation of this method). Diachronic collostruction analysis has been suggested by Hilpert (2011, 2012) with the aim of investigating changes in the semantics of a construction: if a construction comes to be used with different collocates, this may reveal that a semantic change is underway. The diachronic collostruction analysis in Hilpert is an adaptation of a *distinctive* collexeme analysis, which was originally designed to compare the lexical collocates of two or more constructions in synchronic data. In its diachronic application, the analysis is used to compare distinctive collexemes of one construction across time and see which collocates are significantly more frequent than expected in one particular period. Our diachronic application is different in two ways: first, we will apply the *covarying* collexeme analysis and second, we will be comparing separate analyses for subsequent time periods, rather than incorporating data for multiple periods into the same analysis. The implications and interpretation of our analysis are also somewhat different in that we will not primarily be tracking semantic changes in the construction overall. Instead, we are interested in the collocational behaviour of specific verbs and intensifiers, as the combinatorial flexibility was shown to have a crucial influence on the productivity of lower-level subschemas in Chapter 4. By measuring the strength of association between different verbs and intensifiers for different time periods, the covarying collexeme analysis can reveal whether certain associations become stronger or weaker. An increase in the strength of the association may hint at increasing conventionalisation of the collocation, whereas associations growing weaker may indicate that the collocation has debonded or that one or both of the individual items have extended their collocational range to new items. Because collostructional analyses are data-intensive methods, we follow Hilpert (2011, 2012) and merge multiple time periods together in larger clusters, based on the time phases that were formed by the VNC analyses in §5.1. Averaging over the output clusters of all four VNC analyses (i.e. on the basis of token frequency for 1830-1995, verb and intensifier type frequencies for 1830-1995 and token frequency for 1930-1995), we distinguish the following four clusters.

Table 5.9. VNC-based periods for the covarying collexeme analysis

	Conflated corpus size	Total tokens	Total verb types	Total intensifier types
1830s-1930s	1,371,570,578	593	69	58
1940s-1960s	904,003,370,	1,388	89	112
1970s-1980s	596,074,116	2,119	138	118
1990s	297,040,183	2,035	131	115

There are some final methodological cautionary remarks. First, we should only compare the *ranks* of the collocations – not the actual measure of collostruction strength – because we are dealing with data sets that are considerably different in size. Additionally, we should remain wary of over-interpretation. The association strength is heavily dependent upon the individual frequencies of the verb and the intensifier that are part of the collocation, as well as the overall frequency of all items in the data set. In that regard, position shifts in the ranking may hint at ongoing changes in the strength of the particular collocation in question, but they may also be explained by more general frequency changes in the construction overall. Moreover, as it is common practice to only study the top collexemes in collostructional analyses, we have limited our analysis to the top twenty of most attracted collexemes. Of course, this cut-off point is to some extent random and one should keep that in mind when interpreting the results. Frequency fluctuations may cause collocations to suddenly show up in the top twenty or drop outside of the line of sight, but this does not necessarily mean that they suddenly appeared or disappeared in/from the data set. Table 5.10 gives the top twenty attracted collexemes for the first two periods (see Appendix V-3 for the full output).

Table 5.10. Side-by-side comparison of the top 20 attracted collexemes in periods 1 and 2 in Delphcorp (overlap highlighted in grey)

Period 1 (1830s to 1930s)					Period 2 (1940s to 1960s)				
Verb	Int	Obs. Freq.	$\Delta P_{V\text{-}TO\text{-}NT}$/ $\Delta P_{INT\text{-}TO\text{-}V}$	Coll. Str.	Verb	Int	Obs. Freq.	$\Delta P_{V\text{-}TO\text{-}NT}$/ $\Delta P_{INT\text{-}TO\text{-}V}$	Coll. Str.
lopen (28)	het vuur uit de sloffen (23)	23	0.82/ 0.99	36.19	lopen (81)	het vuur uit de sloffen (54)	44	0.54/ 0.79	50.08
denken (30)	suf (67)	28	0.86/ 0.41	26.53	schrikken (304)	een hoedje (63)	57	0.18/ 0.72	32.22
peinzen (24)	suf (67)	22	0.84/ 0.32	19.98	peinzen (34)	suf (87)	27	0.75/ 0.3	27.73
kijken (8)	de ogen uit het hoofd (23)	8	0.97/ 0.35	11.87	zich schamen (81)	dood (374)	68	0.61/ 0.17	27.72
lachen (127)	ziek (26)	22	0.16/ 0.66	11.55	zich ergeren (89)	groen en geel (18)	18	0.2/ 0.95	22.23

lachen (127)	slap (16)	16	0.13/ 0.81	11.05	drinken (12)	een stuk in de kraag (11)	10	0.83/ 0.91	21.99
zich ergeren (58)	dood (215)	44	0.44/ 0.17	9.96	betalen (17)	blauw (38)	14	0.81/ 0.37	20.23
zich schamen (38)	de ogen uit het hoofd (23)	12	0.3/ 0.48	9.19	piekeren (21)	suf (87)	18	0.81/ 0.2	19.36
schreeuwen (9)	schor (17)	6	0.65/ 0.35	7.77	werken (104)	in het zweet (15)	15	0.14/ 0.94	17.31
schrijven (7)	de vingers blauw (5)	4	0.57/ 0.79	7.47	lachen (318)	slap (25)	25	0.08/ 0.79	16.33
zich vervelen (33)	dood (215)	26	0.45/ 0.1	6.45	zich vervelen (79)	dood (374)	55	0.45/ 0.12	15.43
wenen (4)	blind (4)	3	0.75/ 0.75	6.34	lachen (318)	een kriek (30)	27	0.08/ 0.69	14.35
schrikken (56)	een ongeluk (37)	14	0.21/ 0.3	5.98	lachen (318)	krom (44)	34	0.1/ 0.56	13.97
schrikken (56)	een aap (7)	7	0.12/ 0.62	5.35	prakkiseren (16)	suf (87)	13	0.76/ 0.15	13.34
juichen (3)	schor (17)	3	0.98/ 0.18	4.71	schrikken (304)	een aap (65)	41	0.11/ 0.43	12.63
drinken (4)	een stuk in de kraag (2)	2	0.5/ 1	4.47	lopen (81)	de benen PREP het lijf (10)	10	0.12/ 0.95	12.58
kniezen (9)	dood (215)	9	0.65/ 0.04	4.01	denken (15)	suf (87)	12	0.75/ 0.14	12.17
piekeren (4)	suf (67)	4	0.89/ 0.06	3.82	tillen (4)	een breuk (6)	4	1/ 0.67	10.01
werken (56)	in het zweet (5)	4	0.07/ 0.71	3.48	lachen (318)	ziek (33)	25	0.07/ 0.54	9.99
schrijven (7)	de vingers krom (3)	2	0.28/ 0.66	3.45	werken (104)	kapot (48)	19	0.16/ 0.33	9.8

The first thing to note is that, like in Chapter 4, the association is rarely symmetric: in the majority of the cases, the ΔP-values indicate that the association is heavily dependent on the limited combinatorial flexibility of one of the two elements. Second, there is a considerable amount of overlap between periods 1 and 2, with ten specific verb-intensifier combinations being returned as strong collocations in both periods (highlighted in the table). Some of these even occupy the same position in the overall ranking, viz. *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' as the number one strongest collocation, *zich suf peinzen* 'to ponder oneself drowsy' in third place and *zich dood vervelen* 'to bore oneself dead' in 11th place. The added value of using the covarying collexeme analysis to track collocations lies in the fact that it compares observed to expected frequencies rather than just looking at the absolute frequency of

co-occurrence. That is, the top attracted collocation is not the most token frequent one per se, but the collocation for which the observed frequency deviates the most from the expected frequency. This also means that an increase or decrease in absolute co-occurrence frequency does not necessarily imply that the association has become stronger or weaker. For *zich ziek lachen* 'to laugh oneself sick' or *zich slap lachen* 'to laugh oneself weak', for instance, the absolute frequency of these collocations has slightly increased, but both have dropped down a couple of spots in the ranking. The general frequency increase of *lachen* 'to laugh', which is more than twice as frequent in period 2 than in period 1, has boosted the expected frequency of co-occurrence, whereas the observed frequency of co-occurrence has increased only slightly. As the discrepancy between observed and expected frequency shrinks, the association grows weaker because the collocation is perceived as less "surprising". Most of the collocations that have disappeared from the top twenty were low in frequency and, therefore, not particularly interesting. The collocation *zich de ogen uit het hoofd schamen* 'to embarrass oneself the eyes out of the head', however, still has 9 occurrences in the second period, but in combination with the frequency increase of *zich schamen* 'to be embarrassed' and the rise of a "new" collocation *zich dood schamen* 'to embarrass oneself dead', it has dropped out of the top twenty, to the 24th place. "New" is put between quotation marks because *zich dood schamen* 'to embarrass oneself dead' already had 22 occurrences in the first period. In addition, there are some *truly* new verb-intensifier combinations which immediately position themselves at the top of the ranking, viz. *zich een hoedje schrikken* 'to startle oneself a little hat', *zich groen en geel ergeren* 'to annoy oneself green and yellow' and *zich blauw betalen* 'to pay oneself blue'. These collocations are all heavily asymmetric towards the intensifier, an intensifier which was not yet (or barely) attested before the 1940s. In Table 5.11, we continue to track the collocations in the next period, viz. the 1970s and the 1980s.

Table 5.11. Side-by-side comparison of the top 20 attracted collexemes in periods 2 and 3 in Delphcorp (overlap highlighted in grey)

Period 2 (1940s to 1960s)					Period 3 (1970s to 1980s)				
Verb	Int	Obs. Freq.	$\Delta P_{V\text{-}TO\text{-}NT}$/ $\Delta P_{INT\text{-}TO\text{-}V}$	Coll. Str.	Verb	Int	Obs. Freq.	$\Delta P_{V\text{-}TO\text{-}NT}$/ $\Delta P_{INT\text{-}TO\text{-}V}$	Coll. Str.
lopen (81)	het vuur uit de sloffen (54)	44	0.54/ 0.79	50.08	lopen (188)	het vuur uit de sloffen (123)	117	0.62/ 0.92	132.7
schrikken (304)	een hoedje (63)	57	0.18/ 0.72	32.22	zich ergeren (218)	groen en geel (68)	68	0.31/ 0.93	71.79
peinzen (34)	suf (87)	27	0.75/ 0.3	27.73	piekeren (66)	suf (164)	57	0.81/ 0.34	57.47
zich schamen (81)	dood (374)	68	0.61/ 0.17	27.72	zich schamen (93)	dood (290)	67	0.61/ 0.22	39.4

zich ergeren (89)	groen en geel (18)	18	0.2/ 0.95	22.23	schrikken (456)	een hoedje (71)	64	0.14/ 0.71	35.81
drinken (12)	een stuk in de kraag (11)	10	0.83/ 0.91	21.99	betalen (21)	blauw (53)	15	0.7/ 0.28	20.27
betalen (17)	blauw (38)	14	0.81/ 0.37	20.23	werken (226)	uit de naad (55)	33	0.13/ 0.51	18.82
piekeren (21)	suf (87)	18	0.81/ 0.2	19.36	werken (226)	in het zweet (17)	17	0.08/ 0.9	16.76
werken (104)	in het zweet (15)	15	0.14/ 0.94	17.31	lachen (260)	slap (18)	18	0.07/ 0.88	16.63
lachen (318)	slap (25)	25	0.08/ 0.79	16.33	lachen (260)	een bult (27)	22	0.08/ 0.7	15.74
zich vervelen (79)	dood (374)	55	0.45/ 0.12	15.43	lopen (188)	de benen PREP het lijf (21)	17	0.09/ 0.73	14.54
lachen (318)	een kriek (30)	27	0.08/ 0.69	14.35	zich vervelen (96)	dood (290)	43	0.33/ 0.12	13.67
lachen (318)	krom (44)	34	0.1/ 0.56	13.97	prakkiseren (23)	suf (164)	16	0.63/ 0.09	12.9
prakkiseren (16)	suf (87)	13	0.76/ 0.15	13.34	peinzen (11)	suf (164)	11	0.93/ 0.07	12.36
schrikken (304)	een aap (65)	41	0.11/ 0.43	12.63	werken (226)	kapot (173)	49	0.15/ 0.19	11.07
lopen (81)	de benen PREP het lijf (10)	10	0.12/ 0.95	12.58	schrikken (456)	lam (32)	25	0.05/ 0.57	11.05
denken (15)	suf (87)	12	0.75/ 0.14	12.17	piekeren (66)	het hoofd suf (8)	7	0.11/ 0.85	9.79
tillen (4)	een breuk (6)	4	1/ 0.67	10.01	zich ergeren (218)	blauw (53)	23	0.09/ 0.34	9.61
lachen (318)	ziek (33)	25	0.07/ 0.54	9.99	tillen (4)	een breuk (10)	4	1/ 0.4	9.6
werken (104)	kapot (48)	19	0.16/ 0.33	9.8	drinken (6)	een stuk in de kraag (6)	4	0.67/ 0.67	9.57

Looking at the highlighted cells in Table 5.11, we find an even greater overlap between periods 2 and 3 than between periods 1 and 2, with 15 shared collocations in the top twenty. Although both *lopen* 'to run' and *het vuur uit de sloffen* 'the fire out of the slippers' look to be extending their use beyond the mutual combination, the collocation *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' continues to hold the top position. Most of the new collocations that were introduced in the previous period have gained a couple of spots in the ranking: *zich groen en geel ergeren* 'to annoy oneself green and yellow' and *zich suf piekeren* 'to worry oneself drowsy' have risen to rank two and three, respectively, pushing down *zich een hoedje schrikken* 'to startle oneself a little hat' to

the fourth position. Again, there are some "new" collocations that were already attested in the previous period but were still outside of the top twenty, viz. *zich lam schrikken* 'to startle oneself lame' (7 attestations in the previous period) and *zich blauw ergeren* 'to annoy oneself blue' (14 attestations in period 2). A more interesting case is *zich uit de naad werken* 'to work oneself out of the seam', which had only 2 total occurrences in the previous period and has really risen to prominence. Finally, some collocations, like *zich ziek lachen* 'to laugh oneself sick', *zich een aap schrikken* 'to startle oneself a monkey' and *zich suf denken* 'to think oneself drowsy' have now disappeared from the top twenty. Although this does not mean that these verbs and intensifiers have ceased to co-occur (cf. supra), we know that *zich een aap schrikken* and *zich ziek lachen* are unattested or highly infrequent in present-day Netherlandic Dutch (see Chapter 4). This shows that, while changes in the top twenty collocates are often no more than inconsequential fluctuations, they sometimes hint at more essential shifts in the collocational patterns. Finally, we compare the top collocations in periods 3 and 4 in Table 5.12.

Table 5.12. Side-by-side comparison of the top 20 attracted collexemes in periods 3 and 4 in Delphcorp (overlap highlighted in grey)

Period 3 (1970s to 1980s)					Period 4 (1990s)				
Verb	Int	Obs. Freq.	$\Delta P_{V\text{-}TO\text{-}NT}$/ $\Delta P_{INT\text{-}TO\text{-}V}$	Coll. Str.	Verb	Int	Obs. Freq.	$\Delta P_{V\text{-}TO\text{-}NT}$/ $\Delta P_{INT\text{-}TO\text{-}V}$	Coll. Str.
lopen (188)	het vuur uit de sloffen (123)	117	0.62/ 0.92	132.7	zich ergeren (270)	groen en geel (107)	107	0.4/ 0.92	103.26
zich ergeren (218)	groen en geel (68)	68	0.31/ 0.93	71.79	lopen (134)	het vuur uit de sloffen (76)	69	0.51/ 0.87	81.16
piekeren (66)	suf (164)	57	0.81/ 0.34	57.47	piekeren (60)	suf (170)	53	0.82/ 0.31	52.39
zich schamen (93)	dood (290)	67	0.61/ 0.22	39.4	schrikken (470)	een hoedje (66)	66	0.14/ 0.79	43.63
schrikken (456)	een hoedje (71)	64	0.14/ 0.71	35.81	zich schamen (106)	dood (243)	67	0.54/ 0.25	38.12
betalen (21)	blauw (53)	15	0.7/ 0.28	20.27	betalen (37)	blauw (51)	24	0.64/ 0.46	31.71
werken (226)	uit de naad (55)	33	0.13/ 0.51	18.82	werken (224)	uit de naad (71)	46	0.19/ 0.56	28.02
werken (226)	in het zweet (17)	17	0.08/ 0.9	16.76	zich ergeren (270)	bont en blauw (32)	30	0.11/ 0.82	24.36
lachen (260)	slap (18)	18	0.07/ 0.88	16.63	schrikken (470)	rot (257)	129	0.19/ 0.31	23.95

lachen (260)	een bult (27)	22	0.08/ 0.7	15.74	zich vervelen (85)	te pletter (136)	38	0.4/ 0.25	23.68
lopen (188)	de benen PREP het lijf (21)	17	0.09/ 0.73	14.54	lopen (134)	de benen PREP het lijf (18)	16	0.12/ 0.83	17.15
zich vervelen (96)	dood (290)	43	0.33/ 0.12	13.67	schreeuwen (15)	de longen uit het lijf (45)	11	0.72/ 0.24	15.67
prakkiseren (23)	suf (164)	16	0.63/ 0.09	12.9	klappen (8)	de handen stuk (6)	6	0.75/ 1	15.54
peinzen (11)	suf (164)	11	0.93/ 0.07	12.36	drinken (12)	een stuk in de kraag (9)	7	0.58/ 0.78	15
werken (226)	kapot (173)	49	0.15/ 0.19	11.07	betalen (37)	scheel (15)	10	0.27/ 0.65	14.52
schrikken (456)	lam (32)	25	0.05/ 0.57	11.05	lachen (166)	een kriek (13)	13	0.08/ 0.92	14.34
piekeren (66)	het hoofd suf (8)	7	0.11/ 0.85	9.79	praten (7)	de blaren op de tong (5)	5	0.71/ 1	13.14
zich ergeren (218)	blauw (53)	23	0.09/ 0.34	9.61	zuipen (20)	klem (9)	6	0.3/ 0.66	10.49
tillen (4)	een breuk (10)	4	1/ 0.4	9.6	werken (224)	in het zweet (17)	13	0.06/ 0.66	9.4
drinken (6)	een stuk in de kraag (6)	4	0.67/ 0.67	9.57	zich schamen (106)	de ogen uit het hoofd (13)	9	0.08/ 0.64	8.92

Aside from some minor position switches, the collocations in the top six are identical in the third and fourth periods. In general, although the overlap between immediately adjacent periods is quite substantial, there are only a couple of collocations that are historically robust in all four periods, viz. *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' and *zich suf piekeren* 'to worry oneself drowsy', which have been in the top twenty since the first period. Several of the newly introduced collocations in period 2 have taken up a strong position in the top five and the collocation *zich uit de naad werken* 'to work oneself out of the seam', which was introduced in the third period, also maintains its sixth position. The results also hint at a change in collocational preferences for *zich vervelen* 'to be bored' from *dood* 'dead' in period 3 to *te pletter* 'to smithereens' in period 4. Given that there are no large discrepancies in the overall size of the data sets and the individual frequencies of the specific verb and intensifiers involved, this is likely a real shift in preferences rather than random frequency fluctuation. Lastly, there are also two "re-entries", viz. *zich een kriek lachen* 'to laugh oneself a hump' and *zich de ogen uit het hoofd schamen* 'to embarrass oneself the eyes out of the head'. Those had

temporarily dropped out of the top twenty collocates in the third period, but even then they still had a significant collostruction strength of 4.57 (rank 38) and 4.51 (rank 40), respectively.

The diachronic application of the covarying collexeme analysis can clearly offer some useful insights, as long as we remain aware of its limitations when interpreting the results. First of all, while §5.1 may have given the impression that – aside from some oscillations – there is a kind of general, construction-wide expansion going on, this section has demonstrated that we should add some nuance to that statement. While it is certainly true that we find an overall increase in the frequency of use of the construction and the different types of verbs and intensifiers that can fill the open slots, there are important low-level interactions and peculiarities that risk being ignored if we were to only focus on the overall development of the construction. For one, when considering the frequency developments of individual verbs and intensifiers, one should also take into account the collocations they are part of. Over time, certain specific verbs and intensifiers may start to co-occur more frequently than would be expected based on their individual frequencies. The covarying collexeme analysis detects the emergence of several verb-intensifier combinations that have conventionalised or even fossilised into strong collocations in present-day Dutch. In some cases, a newly introduced verb or intensifier pairs up with an intensifier or verb that already had a certain collocational range in the construction, thus immediately forming an asymmetric association. If the new slot filler continues to be (nearly) exclusively attached to the more flexible item, this verb-intensifier combination further develops as an asymmetric collocation, the strength of which is primarily determined by the limited combinatorial flexibility of one of the two elements. With hindsight, several of the top verbs and intensifiers in Figure 5.9 and Figure 5.13 have only increased their frequency by virtue of being part of a specific collocation. In other cases, a collocation appears to have originally been introduced as a more or less symmetric combination, with both verb and intensifier showing strong mutual attraction. Some of these continue to display symmetry (e.g. *zich een stuk in de kraag drinken* 'to drink oneself a piece in the collar', *zich schor schreeuwen* 'to scream oneself hoarse') but others have shifted to a more asymmetric collocation as one of the two items emancipated itself from the other. In case of *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers', the early developments of verb and intensifier are tightly intertwined because they both increased in frequency as part of the same collocation. If we look at the ΔP-values in the tables above, we find that *zich het vuur uit de sloffen lopen* 'the fire out of the slippers' has developed from an almost perfectly symmetric collocation (high values for both int-to-verb and verb-to-int) to a heavily intensifier-asymmetric collocation (high value for int-to-verb, lower value for verb-to-int), which suggests that the verb *lopen* 'to run' has emancipated itself from the collocation to a greater extent than the intensifier. The covarying collexeme analysis did not give any obvious examples of the opposite development, in which an asymmetric collocation gradually *gains* symmetry

– which could be indicative of a narrowing collocational range of the erstwhile more flexible item – but it is definitely a theoretically plausible scenario that needs to be kept in mind as we continue exploring the collocational behaviour in the next sections. Of course, there are also many examples of verbs and intensifiers which appear to have developed "independently" of any specific combinations. Even so, the frequency graphs in §5.1 already showed that not all of these have followed the same trajectory, so we can expect to find differences in the development of their collocational behaviour as well.

While the covarying collexeme analysis allows for a first exploration of some collocational patterns in the construction, it only provides a snapshot of the collocational behaviour and the changes in the collocational range of specific verbs and intensifiers. For example, we find that *lachen* 'to laugh' appears to be part of multiple strong combinations, some of which seem to come and go at random (e.g. *zich krom lachen, zich een kriek lachen, zich een bult lachen, zich ziek lachen, zich slap lachen*...). However, as these collocations do not even account for half of the attestations of *lachen* overall, they can only tell us so much about the collocational behaviour of *lachen*. It is maintained that the covarying collexeme analysis does not take into account certain frequency aspects that are important to the investigation of collocational patterns, i.e. type frequency, hapax count and relative entropy (see Chapter 4 for the cross-tabulation approach). As most of these also play a crucial role in the frequency-based measures of productivity, the issues of collocational expansion and/or conventionalisation will be further elaborated in the next section.

5.3 Shifts in productivity

In Chapter 4, we illustrated that the use of the construction is characterised by both conventionalised highly frequent collocations and infrequent, creative coinages. The results in the previous two sections suggest that this interplay between conventionalisation and productivity has been an important aspect of the development of the construction for several decades. In §5.1 we showed that, while many new verbs and intensifiers have joined the distribution of the construction, some were more successful than others. At the same time, there is a remarkable parallelism in the frequency developments of both verbs and intensifiers. Given the interactions between the two open slots of the construction, it is not always easy to disentangle the development of individual verbs and intensifiers (cf. §5.2). In this section, we discuss the implications of these developments on the productivity of schemas and subschemas at different levels in the constructional network hierarchy. While the overall increase of attested verbs and intensifier types suggests that the construction has without doubt

become more productive at the most abstract level [SUBJ V REFL INT] over time, the observed idiosyncratic changes for specific verbs and intensifiers indicate that there are also subtler, lower-level shifts taking place. We analyse the data within the multidimensional model of productivity that was set out in Chapters 2 and 4.

5.3.1 A frequency-based productivity complex

Based on the observation in §5.1 that the late 20th Century is characterised by more lexical variation and creativity compared to the 19th Century, we might expect to find an overall increase in productivity at the level of the macro-schema [SUBJ V REFL INT]. Consider the frequency-based measures in Table 5.13.

Table 5.13. Frequency-based productivity measures at the macro-level

Decennium	N	V	n1	$\mathscr{P}$	Hapax/ type	V	n1	$\mathscr{P}$	Hapax/ type
		Intensifier				Verb			
1830s	6	4	3	0.50	0.75	5	4	0.67	0.80
1850s	9	8	8	0.89	1.00	6	4	0.44	0.67
1870s	56	15	8	0.14	0.53	17	8	0.14	0.47
1890s	96	23	9	0.09	0.39	25	9	0.09	0.36
1910s	157	29	14	0.09	0.48	38	20	0.13	0.53
1930s	271	37	19	0.07	0.51	39	16	0.06	0.41
1940s	316	57	22	0.07	0.39	42	22	0.07	0.52
1950s	574	71	35	0.06	0.49	53	21	0.04	0.40
1960s	498	65	25	0.05	0.38	56	24	0.05	0.43
1970s	842	97	43	0.05	0.44	96	53	0.06	0.55
1980s	1,277	76	21	0.02	0.28	94	46	0.04	0.49
1990s	2,035	115	39	0.02	0.34	131	65	0.03	0.50

In terms of realised productivity, measured by the number of different types, both the verb and intensifier slots at the maximum level of abstraction appear to have increased their productivity (with some ups and downs in the second half of the 20th Century, cf. supra). Between the 1890s and the 1990s both slots expanded their range of attested fillers from around twenty types to over a hundred different types. The same cannot be said for the potential productivity measure: both slots have undergone a gradual, rather consistent decrease of the $\mathscr{P}$-score (also compare to the $\mathscr{P}$-score of 0.02 in SoNaR-NL). It would appear as if there was more "growth potential" (potential productivity) when the construction was still fairly infrequent and showed little lexical variation, compared to a time when there are already 115 different intensifier types and 131 verb types. When the total number of tokens increases, the type frequency and hapax count increase as well, but they do not do so at the same rate as the token frequency (Baayen & Lieber 1991: 811). Compared to the 1890s, the token frequency is multiplied by 21, whereas the verb

and intensifier hapax counts increase by 8 and 4, respectively. It was mentioned that the exponential frequency increase in this construction was found to be carried by just a small number of highly frequent verbs and intensifiers (cf. supra). If we want to take the high token frequency items that deflate the $\mathscr{P}$-score out of the equation, we can look at the proportion of hapaxes to all *types* instead. The overall decrease is less noticeable in the hapax-type ratio of the intensifier slot, although the number of hapax legomena does appear to be dwindling in the late 20th Century (and in present-day Dutch, cf. the hapax-type ratio of 0.34 in SoNaR-NL). For the verb slot, about half of all verb types are one-offs in all decennia. As we know that all frequency-based measures in general are highly sensitive to token frequency (cf. Chapter 4 and the references therein), we recalculated the frequency-based measures for the largest shared sample size. For the individual decennia, the largest shared sample size is evidently much too small to do any serious analyses, but we can use the merged time periods that were formed on the basis of the VNC analysis in the previous section. Concretely, we extracted a random sample of 593 tokens (the largest shared sample size) for each period, see Table 5.14.

Table 5.14. Frequency-based productivity measures for four sample sets (N=593)

Period	N	V	N_1	$\mathscr{P}$	Hapax/ type	V	N_1	$\mathscr{P}$	Hapax/ type
		Intensifier				Verb			
Period 1	593	59	25	0.042	0.424	69	36	0.061	0.522
Period 2	593	74	31	0.052	0.419	58	27	0.046	0.466
Period 3	593	73	32	0.054	0.438	73	38	0.064	0.521
Period 4	593	76	32	0.054	0.421	65	33	0.056	0.508

While the entire Delphcorp data set is characterised by a gradual increase in absolute frequency of tokens, types and hapax legomena for both verbs and intensifiers, the equally sized subsets tell a slightly different story. The verb slot no longer shows a clear increase in type frequency or hapax count. Clearly, the construction has not exhausted its range of potential verb slot fillers yet when only 593 tokens are sampled: for instance, only about half of the total verb types and hapaxes (65/131 types and 33/65 hapaxes) are represented in the subset of period 4. For the intensifier slot, on the other hand, the overall expansion is also somewhat visible in the subsets, the main increase being situated between periods 1 and 2 (59 to 74 intensifier types and 25 to 31 hapax types). In general, the proportion of the total intensifier types and hapaxes represented in the subsets is higher for the intensifiers than for the verbs (e.g. 76/115 intensifiers and 33/39 hapaxes for period 4). In other words, as the sample size increases, we are more likely to come across new verbs than new intensifiers. As this is exactly what is supposed to be captured by the potential productivity, we would expect a higher $\mathscr{P}$-score for the verb slot, but the difference is extremely small. We will return to the interpretation of the $\mathscr{P}$-score below. The comparison of Table 5.13 and Table 5.14 is primarily interesting because it

appears that the proportion of hapax legomena to the total of all tokens and types actually stays fairly constant across time within the smaller subsets, both for verbs and intensifiers. This suggests that the decrease in ratios in Table 5.13 was indeed heavily influenced by the general increase in token frequency (of a number of highly frequent verbs and intensifiers) and that the creative potential of the construction – to the extent that we can measure this by looking at the hapaxes –, has not changed all that much since the 1930s.

Although it is definitely worthwhile to look at the overall development of the construction at the macro-level, Chapter 4 demonstrated that we need to zoom in on lower-level subschemas in which the verb or the intensifier are specified if we want to get a grasp of the sophisticated structure of the constructional network (see also Ch6, §6.2.2 and §6.2.3 on studying productivity at different levels of abstraction). Ideally, we should also compare the different verbs and intensifiers at their largest shared sample size and extract equally sized subsets for all items in all periods, but as this is not possible due to the low frequencies of individual verbs and intensifiers, we will work with the full data set from now on.

5.3.1.1 Intensifiers

As we know from the previous section, the majority of the intensifiers display an overall upward trend in absolute frequency, the exceptions being *een aap* 'a monkey' and to a lesser extent *krom* 'bent'. Additionally, some of the more recently introduced intensifiers like *rot* 'rotten', *kapot* 'broken' and *te pletter* 'to smithereens' have quickly become dominant. This is also obvious in Table 5.15, which presents the frequency information (the normalised token frequencies are between brackets) and frequency-based productivity measures for the top 15 intensifiers in the entire data set.

Table 5.15. Productivity development of the top 15 intensifiers in Delphcorp

Intensifier	Period	$N_{absolute}$	$N_{relative}$	V	n1	$\mathscr{P}$	Hapax/type
dood							
	Period 1	215 (1.57)	0.36	25	14	0.07	0.56
	Period 2	374 (4.14)	0.27	24	13	0.03	0.54
	Period 3	290 (4.87)	0.14	23	13	0.04	0.57
	Period 4	243 (8.18)	0.12	17	8	0.03	0.47
rot							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	41 (0.45)	0.03	9	3	0.07	0.33
	Period 3	301 (5.05)	0.14	35	18	0.06	0.51
	Period 4	257 (8.65)	0.13	25	13	0.05	0.52

suf							
	Period 1	67 (0.49)	0.11	10	4	0.06	0.4
	Period 2	87 (0.96)	0.06	13	5	0.06	0.38
	Period 3	164 (2.75)	0.08	49	33	0.2	0.67
	Period 4	170 (5.72)	0.08	53	35	0.21	0.66
kapot							
	Period 1	6 (0.04)	0.01	5	4	0.67	0.8
	Period 2	48 (0.53)	0.03	17	11	0.23	0.65
	Period 3	173 (2.9)	0.08	29	16	0.09	0.55
	Period 4	137 (4.61)	0.07	21	7	0.05	0.33
een ongeluk							
	Period 1	37 (0.27)	0.06	10	4	0.11	0.4
	Period 2	101 (1.12)	0.07	24	16	0.16	0.67
	Period 3	67 (1.12)	0.03	17	10	0.15	0.59
	Period 4	65 (2.19)	0.03	20	14	0.22	0.7
het vuur uit de sloffen							
	Period 1	23 (0.17)	0.04	1	0	0	0
	Period 2	63 (0.7)	0.05	9	6	0.1	0.67
	Period 3	123 (2.06)	0.06	6	4	0.03	0.67
	Period 4	76 (2.56)	0.04	6	4	0.05	0.67
te pletter							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	5 (0.06)	0	4	3	0.6	0.75
	Period 3	102 (1.71)	0.05	25	13	0.13	0.52
	Period 4	136 (4.58)	0.07	24	14	0.1	0.58
een hoedje							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	63 (0.7)	0.05	4	2	0.03	0.5
	Period 3	71 (1.19)	0.03	4	2	0.03	0.5
	Period 4	66 (2.22)	0.03	1	0	0	0
groen en geel							
	Period 1	3 (0.02)	0.01	2	1	0.33	0.5
	Period 2	18 (0.2)	0.01	1	0	0	0
	Period 3	68 (1.14)	0.03	1	0	0	0
	Period 4	107 (3.6)	0.05	1	0	0	0
blauw							
	Period 1	5 (0.04)	0.01	4	3	0.6	0.75
	Period 2	38 (0.42)	0.03	10	7	0.18	0.7
	Period 3	53 (0.89)	0.03	14	11	0.21	0.79
	Period 4	51 (1.72)	0.03	7	5	0.1	0.71

uit de naad							
	Period 1	1 (0.01)	0	1	1	1	1
	Period 2	4 (0.04)	0	4	1	0.25	0.25
	Period 3	49 (0.82)	0.02	3	2	0.04	0.67
	Period 4	71 (2.39)	0.03	13	9	0.13	0.69
wezenloos							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	3 (0.03)	0	3	3	1	1
	Period 3	61 (1.02)	0.03	26	18	0.3	0.69
	Period 4	66 (2.22)	0.03	23	17	0.26	0.74
een aap							
	Period 1	10 (0.07)	0.02	2	0	0	0
	Period 2	65 (0.72)	0.05	4	1	0.02	0.25
	Period 3	18 (0.3)	0.01	2	0	0	0
	Period 4	8 (0.27)	0	5	2	0.25	0.4
wild							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	15 (0.17)	0.01	9	7	0.47	0.78
	Period 3	49 (0.82)	0.02	11	7	0.14	0.64
	Period 4	35 (1.18)	0.02	2	0	0	0
krom							
	Period 1	19 (0.14)	0.03	3	1	0.05	0.33
	Period 2	44 (0.49)	0.03	6	4	0.09	0.67
	Period 3	21 (0.35)	0.01	5	3	0.14	0.6
	Period 4	12 (0.4)	0.01	3	1	0.08	0.33

A new measure in Table 5.15 is the relative frequency. Measuring the proportion of the data that is accounted for by this particular intensifier gives an idea of the prominence of the intensifier in question in that particular period. This measure is perhaps not so much a direct measure of productivity per se, but it gives us important insight in the competition among different intensifiers, which is expected to increase as new types are continuously being added to the repertoire of slot fillers (see Ch2, §2.3.2 on the "fevered competition" between intensifiers). While most intensifiers have remained quite stable with respect to their prominence in the construction, there are some indications of a power struggle, especially among the top five intensifiers (see Ch6, §6.2.2 for some discussion on the possible outcomes of competition). The obvious "victim" of this competition is *dood* 'dead': while it has continued to increase in absolute frequency and is still one of the most prominent intensifiers in the most recent period, it has gone from accounting for over a third of all data to "only" 12%. With respect to the potential and realised productivity, as well, *dood* appears to be losing ground. Although it is still a popular intensifier when it comes to boosting the more frequent verb types (i.e. *zich ergeren* 'to be annoyed', *lachen* 'to laugh', *zich vervelen* 'to be annoyed', *schrikken* 'to be startled', etc.), it is attracting less hapaxes compared to earlier stages of Dutch. A less

substantial decrease in relative frequency is attested for *een ongeluk* 'an accident' and *suf* 'drowsy', although both intensifiers show an increase in the other productivity measures. It appears that *suf* 'drowsy' has become the preferred intensifier for infrequent verb types, given the high hapax counts. The intensifiers *rot* 'rotten' and *kapot* 'broken' seem to have come out ahead as "winners" in period 3, judging by an increase on almost all fronts (relative frequency, type frequency and hapax count), but they are slightly less successful in the most recent period. Of course, it needs to be pointed out that while the top five intensifiers seem to be competing amongst one another for the top spot, all of them are also in competition with many infrequent intensifiers. Although it may appear that way by looking at Table 5.15, the decline of *dood* 'dead' should not be interpreted as due to simple replacement by *rot* 'rotten' or *kapot* 'broken'. The top intensifiers risk becoming so frequent that they are no longer felt to be sufficiently expressive or "extravagant" for specific purposes. In usage contexts that require a more expressive intensifier, the frequent "neutral" intensifiers are replaced by a new, infrequent intensifier rather than by another "conventional" or frequent intensifier.

The exclusive association of *groen en geel* 'green and yellow' with *zich ergeren* 'to be annoyed' is evident in the zero scores for most productivity measures – aside from the one curious example with *zich vervelen* 'to be bored' in the first period (cf. supra). In present-day Dutch, the intensifier *het vuur uit de sloffen* 'the fire out of the slippers' also enters into an exclusive (i.e. non-productive) combination with *lopen* 'to run' (cf. Chapter 4), but the diachronic data show that there was some variation in the verbs occurring with this intensifier in the past. The nature of this variation will be further investigated in §5.3.2 below, where we will take a look at the *kinds* of verbs that have appeared with some intensifiers.

The interpretation of the $\mathscr{P}$-score in diachronic studies of productivity is not necessarily straightforward. In Chapter 2 it was mentioned that "the likelihood of being extended to a new type as the sample size increases" could to some extent also be interpreted as "the likelihood of being extended to a new type in the (very near or immediate) future". It is interesting to note that some intensifiers with a high $\mathscr{P}$-score and hapax-type ratio in one period (often the period in which they were first introduced) do in fact gain a considerable amount of types and hapaxes in the next period. *Wezenloos* 'vacant/blank' is the best case in point. With both ratios at the maximum value of 1 in period 2, the type frequency and hapax count are increased by 23 and 15, respectively, in period 3. Another example is *te pletter* 'to smithereens', with a $\mathscr{P}$-score of 0.60 and a hapax-type ratio of 0.75 in period 2, gaining 21 verb types and 10 hapaxes by period 3. K*apot* 'broken', as well, gains 12 types and 7 hapaxes in period 2 after having a potential productivity score of 0.67 and hapax-type ratio of 0.80 in period 1. Given the sensitivity of the $\mathscr{P}$-measure to high token frequency, the hapax-type ratio is sometimes the more reliable measure. The intensifier *uit de naad* 'out of the seam' for example, has a very low

$\mathscr{P}$-score of 0.04 in period 3, but the hapax-type ratio of 0.67 more accurately highlights the proportion of hapaxes. By the same logic, the fact that some of the highly frequent intensifiers like *dood* 'dead', *rot* 'rotten' and *kapot* 'broken' have lower ratio scores in the most recent period may indicate that they have nearly exhausted their productive potential and are less likely to still be extended to new (infrequent) verb types – indeed, their type and hapax counts have stayed relatively constant between the 1990s and present-day Dutch (cf. Chapter 4). Even so, high ratio scores are not a necessary prerequisite for collocational expansion, as is demonstrated by *rot* 'rotten'. In spite of the rather low ratios in period 2, there is a considerable increase in both type frequency (9 to 35 types) and hapax count (3 to 18) between periods 2 and 3. What is more, the intensifier may *decrease* in productivity despite having high ratios in the previous period. This is the case for *wild* 'wild', which has a hapax-token ratio of 0.14 and hapax-type ratio of 0.64 in period 3 but still drops down from 11 to 2 types and 7 to 0 hapaxes. The same could be said for *een aap* 'a monkey', which still has a hapax-token ratio of 0.25 and a hapax-type ratio of 0.40 in the final period of Delphcorp but has completely disappeared in present-day Netherlandic Dutch. We will return to the issue of using hapax-token and hapax-type ratios in historical productivity in Chapter 6, §6.2.2.

In order to quickly gauge the changes in productivity of the intensifiers over time, we can plot the ($\mathscr{P}$,V) coordinates for the different periods in the global productivity graph. Generally speaking, a movement towards the top right indicates an overall increase in global productivity, whereas a shift towards the bottom left of the plane is interpreted as a decrease in global productivity. Adding four data points for twenty intensifiers would result in a very cluttered graph, so only the top five was plotted in Figure 5.15.

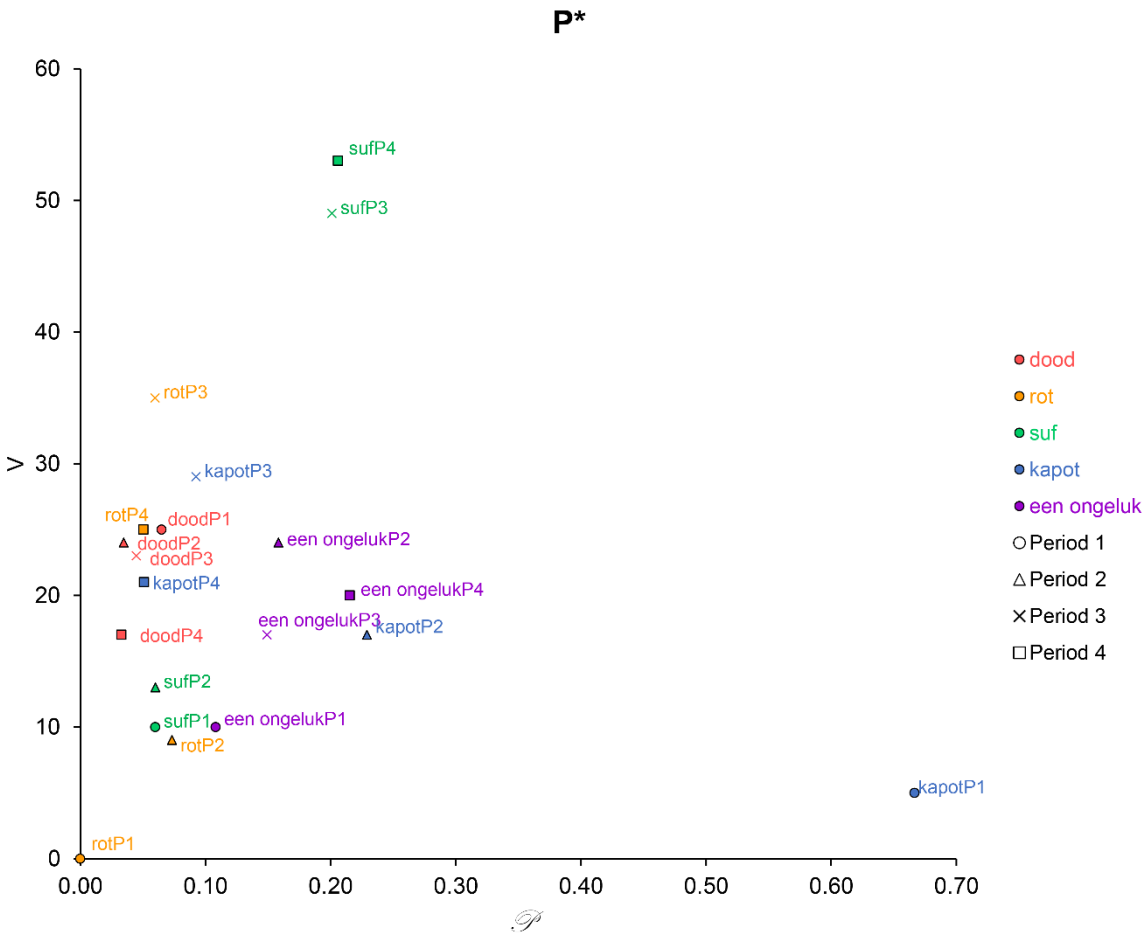


Figure 5.15. Global productivity for the top 5 intensifiers in Delphcorp

If we were to just compare the first period to the most recent period, we could state that *rot* 'rotten', *een ongeluk* 'an accident' and *suf* 'drowsy' show an overall increase in global productivity, with both productivity coordinates being higher in P4 than in P1, whereas *dood* 'dead' has decreased with respect to both aspects of productivity. The development of *kapot* 'broken' cannot be summarised as more or less (globally) productive because it is characterised by an increase in realised productivity but a decrease in potential productivity. Taking into account the intermediary periods, we find that the development of most intensifiers is not exactly linear. Both *een ongeluk* 'an accident' and *rot* 'rotten' display some fluctuation in potential productivity, as well as realised productivity. Whereas *kapot* 'broken' consistently slips to the left while moving up and down the Y-axis, *dood* 'dead' is consistently shifting downwards, while showing some horizontal fluctuations. In addition to tracking the development for each individual intensifier, the global productivity graph allows us to identify intensifiers which either appear to be following similar paths or, conversely, embark on very different pathways. For example, whereas *rot* 'rotten' and *kapot* 'broken' were on opposite sides of the X-axis in the first period, they have become relatively close together by the second half of the 20th Century (periods 3 and 4). Conversely, *suf* 'drowsy' and *een ongeluk* 'an accident' were

in each other's proximity until the 1930s (period 1), but *suf* has ended up with a much higher type frequency by the end of the 20th Century. The graph also reveals differences with respect to the speed at which the intensifiers develop. Whereas *kapot* 'broken' displays a large gap between P1 and P2, the breakthrough of *suf* 'drowsy' is situated between P2 and P3. In contrast, all four data points for *dood* 'dead' are clustered quite closely together, indicating that its degree of productivity has all in all not changed that much (see §5.4 for the implications of this finding on the position of [SUBJ V REFL *dood*] in the constructional network). In what follows, we switch from the perspective of the intensifier to the perspective of the verb.

5.3.1.2 Verbs

Table 5.16 provides an overview of the relevant frequency measures for the top twenty verbs in Delphcorp.

Table 5.16. Productivity development of the top 15 verbs in Delphcorp

VERB	PERIOD	$N_{ABSOLUTE}$	$N_{RELATIVE}$	V	N1	$\mathscr{P}$	HAPAX/TYPE
schrikken							
	Period 1	56 (0.41)	0.09	12	8	0.14	0.67
	Period 2	304 (3.36)	0.22	31	13	0.04	0.42
	Period 3	456 (7.65)	0.22	35	12	0.03	0.34
	Period 4	470 (15.82)	0.23	32	15	0.03	0.47
lachen							
	Period 1	127 (0.93)	0.21	21	11	0.09	0.52
	Period 2	318 (3.52)	0.23	35	11	0.03	0.31
	Period 3	260 (4.36)	0.12	35	12	0.05	0.34
	Period 4	166 (5.59)	0.08	29	13	0.08	0.45
zich ergeren							
	Period 1	58 (0.42)	0.1	8	3	0.05	0.38
	Period 2	89 (0.98)	0.06	18	10	0.11	0.56
	Period 3	218 (3.66)	0.1	17	5	0.02	0.29
	Period 4	270 (9.09)	0.13	25	12	0.04	0.48
werken							
	Period 1	56 (0.41)	0.09	11	4	0.07	0.36
	Period 2	104 (1.15)	0.07	25	14	0.13	0.56
	Period 3	226 (3.79)	0.11	32	10	0.04	0.31
	Period 4	224 (7.54)	0.11	39	16	0.07	0.41
lopen							
	Period 1	28 (0.2)	0.05	7	3	0.11	0.43
	Period 2	81 (0.9)	0.06	19	13	0.16	0.68
	Period 3	188 (3.15)	0.09	30	15	0.08	0.5
	Period 4	134 (4.51)	0.07	23	10	0.07	0.43

zich schamen							
	Period 1	38 (0.28)	0.06	4	0	0	0
	Period 2	81 (0.9)	0.06	7	5	0.06	0.71
	Period 3	93 (1.56)	0.04	6	1	0.01	0.17
	Period 4	106 (3.57)	0.05	7	3	0.03	0.43
zich vervelen							
	Period 1	33 (0.24)	0.06	6	3	0.09	0.5
	Period 2	79 (0.87)	0.06	13	7	0.09	0.54
	Period 3	96 (1.61)	0.05	11	7	0.07	0.64
	Period 4	85 (2.86)	0.04	8	3	0.04	0.38
piekeren							
	Period 1	4 (0.03)	0.01	1	0	0	0
	Period 2	21 (0.23)	0.02	4	3	0.14	0.75
	Period 3	66 (1.11)	0.03	4	2	0.03	0.5
	Period 4	60 (2.02)	0.03	4	2	0.03	0.5
zoeken							
	Period 1	9 (0.07)	0.02	6	4	0.44	0.67
	Period 2	22 (0.24)	0.02	14	11	0.5	0.79
	Period 3	45 (0.75)	0.02	12	4	0.09	0.33
	Period 4	27 (0.91)	0.01	12	8	0.3	0.67
trainen							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	4 (0.04)	0	3	2	0.5	0.67
	Period 3	32 (0.54)	0.02	15	10	0.31	0.67
	Period 4	45 (1.51)	0.02	16	8	0.18	0.5
peinzen							
	Period 1	24 (0.17)	0.04	3	2	0.08	0.67
	Period 2	34 (0.38)	0.02	4	1	0.03	0.25
	Period 3	11 (0.18)	0.01	1	0	0	0
	Period 4	9 (0.3)	0	2	1	0.11	0.5
betalen							
	Period 1	2 (0.01)	0	2	2	1	1
	Period 2	17 (0.19)	0.01	3	1	0.06	0.33
	Period 3	21 (0.35)	0.01	7	6	0.29	0.86
	Period 4	37 (1.25)	0.02	4	1	0.03	0.25
prakkiseren							
	Period 1	8 (0.06)	0.01	4	3	0.38	0.75
	Period 2	16 (0.18)	0.01	3	1	0.06	0.33
	Period 3	23 (0.39)	0.01	5	2	0.09	0.4
	Period 4	11 (0.37)	0.01	4	3	0.27	0.75

sjouwen							
	Period 1	2 (0.01)	0	2	2	1	1
	Period 2	12 (0.13)	0.01	6	4	0.33	0.67
	Period 3	27 (0.45)	0.01	15	8	0.3	0.53
	Period 4	14 (0.47)	0.01	9	7	0.5	0.78
rijden							
	Period 1	0 (0)	0	0	0	0	0
	Period 2	6 (0.07)	0	5	4	0.67	0.8
	Period 3	22 (0.37)	0.01	15	12	0.55	0.8
	Period 4	26 (0.88)	0.01	12	5	0.19	0.42

If we look at the normalised frequencies, almost all verbs have gradually increased their frequency of occurrence in the construction, but the discrepancy between the first and last periods is much greater for some verbs than for others. The most substantial increases are attested for *schrikken* 'to be startled', *zich ergeren* 'to be annoyed' and *werken* 'to work'. Judging by the relative frequencies, *schrikken* 'to be startled' has clearly risen to prominence in the recent periods: the verb has accounted for over 20% of all examples of the construction in the data set since the mid-20th Century and also started co-occurring with many new intensifier types. In contrast, the position of the most prominent verb in the construction in the first period, viz. *lachen* 'to laugh', has weakened relative to other verbs, although its overall frequency and collocational range are still considerable. The verbs are not in direct competition in the same way as the intensifiers are, because the motivation behind verb selection is evidently different from intensifier selection. That is, while all intensifiers have the same effect of boosting the verbal activity – although they may differ in strength or expressive force – the verbs generally denote different verbal activities and are therefore not interchangeable like the intensifiers. Concretely, if speakers want to express that they are very startled/bored/embarrassed/etc., they have a wide array of intensifiers at their disposal to do so. The reverse scenario, in which a speaker first decides on a specific intensifier and then chooses which activity to combine it with, is highly implausible in natural language use. Nevertheless, a different kind of competition could provide an explanation for the rise and fall of some verbs, viz. a competition at the level of the (schematic) *construction*. The relative decrease of *lachen* 'to laugh' could indicate that language users have come to use a different intensifying pattern to express that they find something very funny. In the same vein, the increase of *schrikken* 'to be startled' could be interpreted as the intensifying fake reflexive resultative construction overtaking another construction as the preferred pattern to express that one is very startled.[60] Potentially competing constructions are the patterns [V *van het* V_{inf}],

[60] Of course, it could also mean that language users are now less likely express that they find something very funny (and more likely to express that they are very startled) than in the 19th Century, but it is not clear why that would be the case.

as in *wenen/brullen/bulderen van het lachen* 'cry/roar with laughter' or [*sterven van* NP] *sterven van schrik/verveling/schaamte* 'die of fright/boredom/embarrassment', which express largely the same meaning as the intensifying fake reflexive resultative construction (cf. Chapter 2). It would be interesting to investigate the overlap in types between such functionally similar constructions and whether there are certain parallels in the development of these constructions.

As we explained in Chapter 4, the potential productivity scores are across the board lower for the verbs than for the intensifiers, so the hapax-type ratio may be a more useful indicator of the extensibility of a verb. Much like in the previous paragraph, we find that high ratios in one period sometimes appear to be announcing a type and hapax increase in the next period. That is the case for, e.g., *schrikken* 'to be startled' and *zoeken* 'to search' between periods 1 and 2 or *lopen* 'to run', *trainen* 'to train' and *sjouwen* 'to drag' between periods 2 and 3. Again, however, not all type or hapax increases are heralded by high ratios in the preceding period. The verbs *zich ergeren* 'to be annoyed' and *werken* 'to work', for example, both gain a lot of types and hapaxes between periods 1 and 2, although less than 10% of all tokens and only about a third of the types were hapaxes in period 1. At the same time, high ratios do not necessarily lead to an increase in the next period, as is demonstrated by, e.g., *piekeren* 'to worry' and *betalen* 'to pay', which have both lost a couple of (hapax) types despite a high $\mathscr{P}$-score and hapax-type ratio in the immediately preceding period. While the hapax counts offer some insight into the expansion of an item, we would have to look at the actual lexical items in the data set to understand exactly *which* new types and hapaxes are being attracted. It is possible that a new hapax is not truly new but recycled from a previous period (cf. the "occasional visitors" in §5.1.3). In addition, it is not because a certain (hapax) type is no longer attested in combination with a verb, that it is no longer a possible collocate of that verb. For example, we see that *piekeren* 'to worry' is found with 4 different intensifier types in periods 2 to 4. A constant value is *suf* 'drowsy', but there is some variation in the other types it combines with. In period 2, the three hapaxes are *gek* 'crazy', *een ongeluk* 'an accident' and *dood* 'dead', in period 3 we find *tureluurs* 'crazy' and *dood* 'dead' again and in period 4 the hapaxes are *rot* 'rotten' and *wezenloos* 'vacant/blank'. Over the entire data set, then, *piekeren* actually occurs with five hapax legomena and there is no reason to assume that, e.g., *zich gek piekeren* 'to worry oneself crazy' is no longer "grammatical" in period 4 (see Ch4, §4.2.1 on why unattested examples are not de facto impossible). We will have a closer look at changes with respect to specific collocates in §5.3.2 below.

The global productivity graph in Figure 5.16 allows us to compare the changes in productivity per individual verb, as well as discover potential trends across different verbs. Parallel to the intensifiers, only the top five verbs were plotted to avoid uninterpretable cluttering.

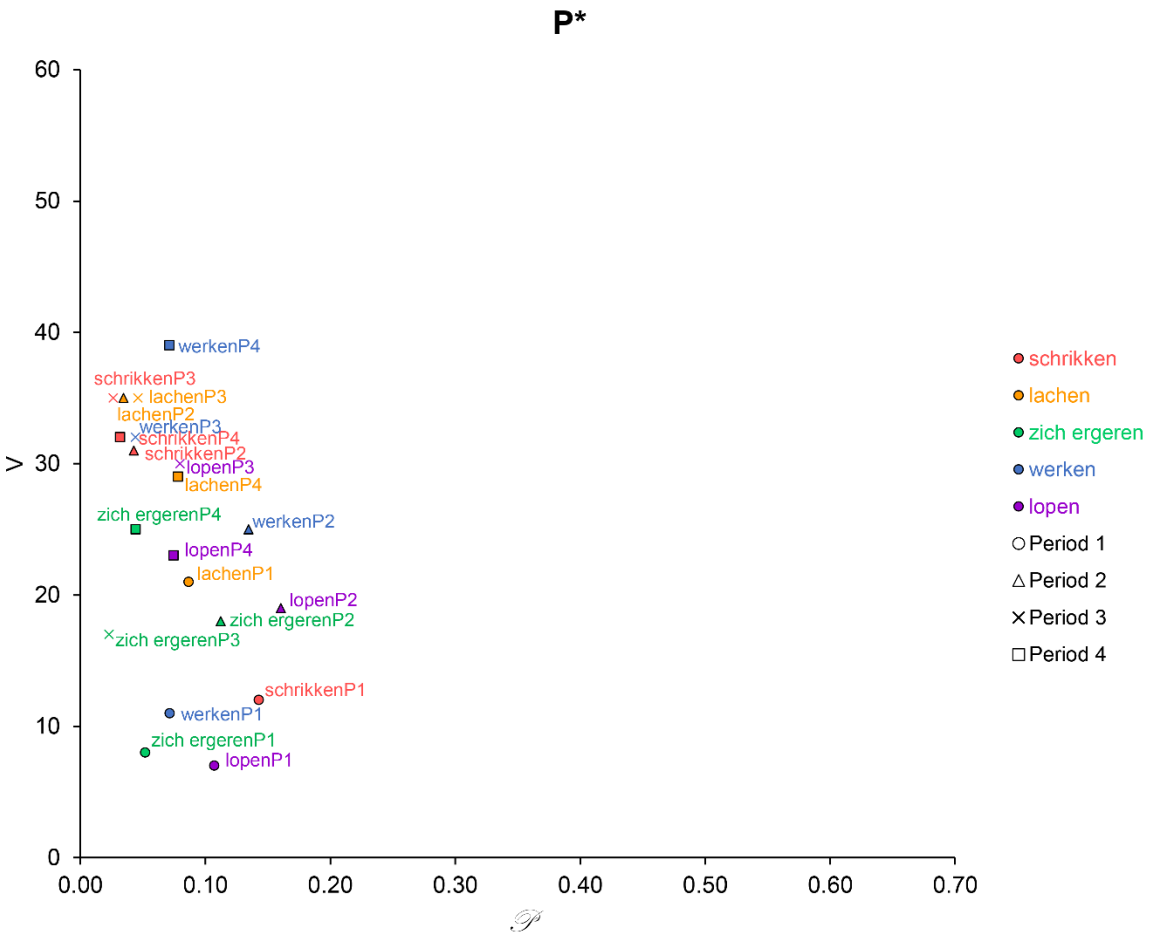


Figure 5.16. Global productivity for the top 5 intensifiers in Delphcorp

Generally speaking, the verbs are slightly less spread out across the plane than the intensifiers. A comparison of the first and last period reveals that none of the verbs display a clear increase or decrease in global productivity. That is, P4 is not situated more to the top right (i.e. increase) or more to the bottom left (i.e. decrease) than P1 for any of the verbs. Instead, while all verbs do move upwards the Y-axis, they are generally shifting slightly to the left on the X-axis. In other words, as the realised productivity increases, the potential productivity shows a slight decrease.

The current section showed that the individual verbs and intensifiers do not only display varying degrees of productivity in present-day Dutch, they also show diverging developments over time. While some verbs and intensifiers have been gradually attracting new collocates, slowly increasing their realised productivity, others much more quickly acquired an elaborate collocational range after they were introduced in the construction. The verbs and intensifiers also differ with respect to their potential productivity, i.e. the likeliness that they are extended to new types. Some items appear to have nearly exhausted their productive potential, while others still have some room to expand their collocational range. However, the frequency-based measures do not tell us anything about the semantic range of specific verbs and intensifiers. As such semantic

aspects were found to be highly relevant to constructional productivity in Chapter 4, the next paragraph will investigate to what extent the type and token expansion of the construction is also a *semantic* expansion.

5.3.2 A constructional model of productivity

As was discussed in Chapters 2 and 4, the relevance of semantics for productivity has long been recognised, but Barðdal (2008) was the first to suggest a theoretical model of productivity that incorporates both frequency and semantic aspects. The underlying assumption of the model is that there is an inverse correlation between type frequency and semantic coherence, as is represented in the by now familiar cline in Figure 5.17, taken from Barðdal (2008: 38).

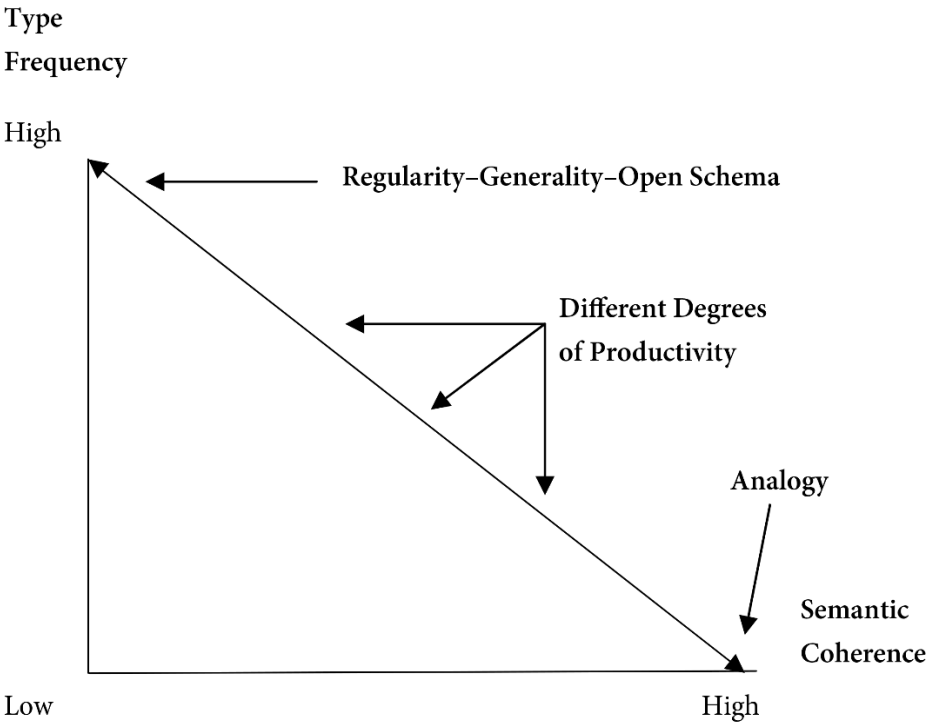


Figure 5.17. Different aspects of the cline of productivity

However, applying the model to the synchronic data in Chapter 4 also revealed that, while there does appear to be a non-trivial interaction between type frequency and semantic coherence, there is some room to refine the model in a number of ways. For the model to become empirically applicable as a measure of productivity, Figure 5.17 could be reinterpreted as a graphical plane in which categories can be plotted at precise (X,Y) coordinates. While we can simply add the type frequency on the Y-axis, we need to find a way to quantify semantic coherence on the X-axis, rather than relying on linguistic intuition or pre-established verb classes. Chapter 4 suggested a number of potential approaches to operationalising semantic coherence, but the best candidate appeared to

be a distributional semantics approach using Vector Space models. These have already been used to study productivity more directly (without direct reference to an inverse correlation with type frequency). The main assumption behind this approach is that changes in the types (of verbs) that are attested in a certain construction may reflect changes in the qualitative productivity of the construction. For example, a detailed look at the recent (i.e. the 1830s to the 2000s) development of the verb distribution of the *way*-construction in Perek (2016a) reveals that the manner-sense, which shows a clear preference for verbs encoding difficulty of motion, gradually opens up its distribution to more neutral motion verbs in the 20th Century. In the path-creation sense, the original semantic domain of literal path-creation verbs continues to grow in the 20th Century, but the construction also becomes more open to abstract ways of creating a (metaphorical) path from the 1880s onwards. In both cases, the semantic constraints on the verb slot appear to have relaxed and, accordingly, the construction has gained in productivity (see also Ch2, §2.1.3). In contrast, Perek (2016b) found that the *hell*-construction (i.e. a construction used to intensify the verb, as in *to scare the hell out of someone*) has expanded its use, but the newly attracted verbs are mostly related to two large semantic domains, the verbs of forceful actions and the verbs of cognition and emotion. In that case, the type frequency increase is not to be interpreted as a semantic expansion, but as an increase of the productivity of the two semantic domains that the construction has been centred on from the beginning. Future research is needed to test how exactly these Vector Space models can be used to quantify semantic coherence in the constructional model of productivity.

If the model in Figure 5.17 were to be operationalised as an empirical tool, some of the theoretical claims of the model may need to be relaxed to some extent. Without question, the main tenets that (i) a schema does not need to rely on semantic coherence if it is sufficiently type frequent and that (ii) a schema with a lower type frequency can still display some degree of productivity if its types display enough semantic coherence, remain valid. In Chapter 4, we found clear examples of intensifiers that combine with a wide range of semantically heterogeneous verbs (following the first tenet), as well as intensifiers that are remarkably productive within a (very) delimited semantic domain (as predicted by the second tenet). However, the inverse correlation, visualised by a straight linear cline between high type frequency and high semantic coherence may be too rigorous. The main issue with this is that, if a category is to be productive (i.e. situated on or close to the cline) in the current model, a lower type frequency necessarily implies a higher degree of semantic coherence, and mutatis mutandis for high type frequency and low coherence. While this is not necessarily a problem for a theoretical model of productivity primarily aimed at showing how type frequency and semantic coherence can interact in robust ways, the current cline may become untenable if we want to actually plot multiple categories (in our case, verbs or intensifiers) on it. In Chapter 4, we gave several examples of intensifiers that all have rather low type frequencies but differ

with respect to their degree of semantic coherence. Within the current model, only the low type frequency intensifiers with high coherence would be situated on the cline, the others being positioned somewhere in the area below the cline. That is, only the categories with high coherence are predicted to be productive (within their delimited semantic domain), whereas the others should hardly show any signs of productivity – but the data suggested otherwise. Conversely, we found several intensifiers occurring with semantically heterogeneous verb types (i.e. low semantic coherence), but which had very different type frequencies. An attempt to plot these somewhere in the graphical plane of the productivity model shows that only the high type frequency intensifiers would be considered productive (i.e. on the cline), but again this prediction was not borne out by our data. One of the questions that was raised in that regard was how to determine the exact threshold at which the type frequency has become "sufficient" so as to no longer "require" semantic coherence, a question that should probably be considered on a construction-by-construction or even item-by-item basis. The main suggestion in Chapter 4 was to abandon the strict linearity of the cline and to approach the model in a more relative way. That is, rather than predicting the degree of productivity of a construction from its position in relation to the cline, we could compare the degree of productivity of multiple constructions by looking at their relative positions to one another, in much the same way as was done for the global productivity graph by Baayen.

Keeping the above reflections in mind, the remainder of this paragraph will not insist too much on the strict inverse correlation, nor will we try to fit specific verbs and intensifiers on the cline. Instead, the focus will be on type expansion and semantic expansion in a more general way, by investigating how the semantic ranges of the verbs and intensifiers have changed over time. In doing so, we subscribe to the same ideas as Perek (2016a, 2016b), viz. that the semantic evolution of a construction may be illuminating in studying productivity shifts.

In §5.1 above, we already mentioned that there is evidence of semantic broadening of both the verb and intensifier slots at the macro-level of the construction [SUBJ V REFL INT]. The verb slot used to be primarily limited to a number of cognitive and emotional verbs like *lachen* 'to laugh', *zich ergeren* 'to be annoyed' or *denken* 'to think', as well as some physical activity verbs like *werken* 'to work' and *lopen* 'to run'. As the verb slot starts to attract more verb types, not only are the existing semantic classes further elaborated, we also see an expansion to verbs from other semantic classes. Over the past two centuries, the construction has come to recruit several verbs of noise emission (e.g. *schreeuwen* 'to scream', *zingen* 'to sing', etc.), verbs of consumption (e.g. *drinken* 'to drink', *eten* 'to eat', etc.) and verbs of communication (e.g. *praten* 'to talk', *discussiëren* 'to discuss', etc.). In addition, from the mid-20th Century onwards, a wide variety of different kinds of activity verbs start showing up in the construction. In present-day Dutch, it appears that virtually every verb that has some inherent aspect that can be boosted, is available for use in the construction. The intensifier slot started out with a limited number of types from the

category of negatively connoted states (e.g. *ziek* 'sick', *dood* 'dead' or *suf* 'drowsy') and the occasional inalienable possession intensifier (e.g. *het vuur uit de sloffen* 'the fire out of the slippers'). As these categories recruited new members over time, it appears that the need for innovation and expressivity has inspired language users to also start looking elsewhere for new intensifiers. The late 19th Century saw the addition of some first terms for bodily ailments (e.g. *een bult* 'a hump', *een koliek* 'a colic', etc.) and the colour terms *blauw* 'blue' and *groen en geel* 'green and yellow' were added to the repertoire in the early 20th Century. From the mid-20th Century onwards, both the category of colour terms and the category of diseases have become very successful. With respect to the colour category, there are several infrequent, creative variations on the frequent models *groen en geel* and *blauw* (cf. supra). In the diseases category, we get a large variety of (often informal) terms for all kinds of both real and fictitious diseases (e.g. *de pleuris* 'the pleurisy', *de tering* 'the consumption', *het schompes* 'fictitious disease', *het leplazerus* 'fictitious disease'...). Around the same time, we start seeing intensifiers that do not easily fit into one of the established semantic categories, including *een slag in de rondte* 'a punch around' and several random objects like *een hoedje* 'a little hat', *een kokosnoot* 'a coconut', *een zuurstok* 'a stick of rock', *een rotje* 'a cracker' or *een mik* 'a loaf'. In sum, the expansion of the construction to new verb and intensifier types is concomitant with a relaxation of the collocational constraints that used to pertain to those constructional slots.

In Chapter 4, we showed that the notion of semantic coherence is mainly relevant if we consider the collocational behaviour at the level of the individual verbs and, especially, intensifiers. In present-day Dutch, there are several examples of original lexical semantics of intensifiers constraining the kinds of verbs they can co-occur with. Although some verbs also show some semantically-motivated collocational preferences (e.g. *lopen* 'to run' with intensifiers involving feet, legs or slippers), the verbs are generally less "picky" about the type of intensifiers they combine with. In the previous subsection, multiple individual verbs and intensifiers were already found to have expanded their collocational ranges in terms of *type frequency*, so this paragraph will investigate whether there is also evidence of a *semantic* expansion. In addition, we will have a closer look at a small set of verbs and intensifiers which appear to be contracting their collocational range.

5.3.2.1 Intensifiers

In present-day Dutch, several intensifiers are found to co-occur with a variety of verbs from different semantic classes, e.g. *dood* 'dead', *te pletter* 'to smithereens', *kapot* 'broken', *een ongeluk* 'an accident', *wezenloos* 'blank/vacant', *een slag in de rondte* 'a punch around'. However, although those intensifiers display a certain flexibility with respect to their combinatorics, that is not to say that they do not show any collocational *preferences* at all: even though they are found to boost a wide variety of verbal activities, some categories

are better represented than others (because they are more prominent in the construction in general). Some intensifiers also have pronounced preferences for one specific verb type, e.g. *wezenloos* 'blank/vacant' and *een ongeluk* 'an accident' both prefer s*chrikken* 'to be startled', even though they are used with a range of other verbs that are not semantically related to that particular verb as well. Neither does the combinatorial flexibility or "all-roundness" mean that the verb slot is *maximally* schematic. The constraint on the verb slot at the highest level of schematicity, i.e. that the verb must have some inherent aspect that is available for intensification (see Ch4, §4.1.1.1), is also inherited by the verb slots in lower-level, partially specified subschemas. In addition, the lower-level schemas may be subject to additional, idiosyncratic constraints. For example, the intensifier *een slag in de rondte* 'a punch around' was found to present itself as an activity booster: while it co-occurs with verbs denoting all kinds of activities (see the examples from Chapter 4), it is not used to intensify any of the emotional or cognitive experience verbs.

If we look at the diachronic development of some of these intensifiers, we find that the semantic expansion of the verb slot at the macro-level is to some extent parallelled in the collocational expansion of individual intensifiers as well. That is, several of the intensifiers appear to have started out with a set of verbs from the two (unrelated) classes of experience verbs and physical activity verbs, which were virtually the only verb classes that were used in the construction overall. The earliest attestations of *dood* 'dead', *kapot* 'broken' and *een ongeluk* 'an accident' in our data demonstrate that they were already used with both major verb classes. Over time, as the verb slot in general became more schematic, the all-round intensifiers expanded their collocational range to more verb types from other semantic classes as well (e.g. *studeren* 'to study', *argumenteren* 'to argue', *zoeken* 'to search'). *Rot* 'rotten' and *te pletter* 'to smithereens' only started appearing around the mid-20th Century, at which point the construction already allowed for several different semantic verb classes. *Rot* 'rotten' was introduced as a hapax with *lachen* 'to laugh' in the 1940s, but it already occurred with other experience verbs and physical activity verbs by the 1950s (e.g. *zich vervelen* 'to be bored', *sjouwen* 'to drag'), as well as several general activity verbs by the 1960s (e.g. *poetsen* 'to clean', *zoeken* 'to search'). The intensifier *te pletter* 'to smithereens' had a rather slow start. It first appeared in the construction in the 1950s with two physical activity verbs, *lopen* 'to run' and *vechten* 'to fight' – which, as was argued in Ch3, §3.2.5.2, may have been a relic of its resultative use – and only has 3 occurrences in the 1960s, viz. two with the experience verb *zich vervelen* 'to be bored' and one with *werken* 'to work'. By the 1970s, it still has only 15 total attestations but it is already occurring with a wide variety of verb classes (e.g. *solliciteren* 'to apply for jobs', *schrijven* 'to write', *roken* 'to smoke', etc.). Although we do see some signs of semantic broadening, the data suggest that several of the present-day all-round intensifiers already displayed some combinatorial flexibility in their early uses as well.

There was no evidence of any idiosyncratic (i.e. intensifier-specific) constraints for the abovementioned intensifiers in the early data in Delphcorp.[61]

(a) Collocational expansion

The data also contain much clearer and more interesting examples of semantic broadening. The intensifier that has without question undergone the most drastic expansion, both in terms of type frequency and in terms of semantic coherence, is *suf* 'drowsy'. As was argued in Chapter 4, the present-day use of *suf* 'drowsy' is characterised by an extremely high degree of combinatorial flexibility, as it is combined with no less than 61 different verb types from different semantic classes. At the same time, in spite of the immense variety of different verb types, there is a remarkably coherent group of cognitive experience and mental activity verbs. We hypothesised that this may well have been the verb class that *suf* 'drowsy' originally started out with, before it started extending its use to all other classes. The diachronic data in Delphcorp now allow us to track its development over the past two centuries. In the 1830s, 1850s and 1870s together, there are 6 examples with *suf* featuring 3 verb types *denken* 'to think' (4), *mijmeren* 'to muse' (1) and *peinzen* 'to ponder' (1).

(236) De goede sukkel krabt achter zijn oor, knijpt in zijn neus, **denkt zich suf** en vraagt bij zichzelven: Wat scheelt mijne vrouw ? (Delphcorp, 1830-1839)
[...] thinks himself drowsy [...]
'The kind fool scratches behind his ear, pinches his nose, thinks hard and asks himself: What is wrong with my wife?'

(237) Dit gewigtig problema, waarover reeds menig schrandere bol **zich suf gemijmerd** heeft [...] (Delphcorp, 1850-1859)
[...] many clever head itself drowsy contemplated has [...]
'This important problem, which many geniuses have contemplated intensely.'

(238) Geen diplomaat heeft er **zich suf** over **gepeinsd** om te verhoeden, dat de Spaansche regeering sedert menschengeheugenis met hare crediteuren naar welgevallen omsprong. (Delphcorp, 1870-1879)
no diplomat has it himself drowsy over pondered [...]
'No diplomat has really thought hard about how to avoid the Spanish government from treating their creditors as they please.'

Until the first half of the 20th Century, *suf* 'drowsy' exclusively co-occurred with verbs that denote some kind of cognitive/mental activity: in addition to *denken* 'to think' and *peinzen* 'to ponder', which account for the majority of the data, we see the verbs *filosoferen* 'to

[61] Of course, it was already mentioned that the first attestation in Delphcorp in all likelihood is not the first occurrence ever of that intensifier in the construction. Especially for *dood* 'dead', we know that it has been at least sporadically since the 17th Century, so we are missing part of its early development.

philosophise', *lezen* 'to read' *prakkiseren* 'to brood', *verzinnen* 'to invent/to think' and, of course, *piekeren* 'to worry'. Based on the strong degree of entrenchment in present-day Dutch of *zich suf piekeren* 'to worry oneself drowsy', we hypothesised in Chapter 4 that *piekeren* 'to worry' may have been the first collocate of *suf* 'drowsy'. We now find that *piekeren* 'to worry' only joined the distribution in the 1910s, but quickly managed to overtake the original collocations by the mid-20th Century. Some examples of the new mental activity verbs are provided in (239) to (241).

(239) Ik **prakkiseerde mij suf**, hoe het toch voor den drommel mogelijk was, dat moeder de vrouw telkens het uur wist, wanneer ik thuis was gekomen. (Delphcorp, 1890-1989)
I brooded myself drowsy [...]
'I've been thinking a lot about how on earth it was possible that my wife always knew at what time I got home.'

(240) Al moet men wonderen van zuinigheid doen en **zich suf piekeren** om er eenigszins dragelijk te komen. (Delphcorp, 1910-1919)
[...] and themselves drowsy worry [...]
'Even if one has to perform miracles to save money and worry a lot about how to survive.'

(241) Ze hadden **zich** al **suf verzonnen**, wat ze hun vriend toch eens konden meegeven als souvenir. (Delphcorp, 1910-1919)
they had themselves already drowsy invented [...]
'They had already thought a lot about what they could give their friend as a souvenir.'

In the 1930s, the mental activity verbs are still largely dominant, but we find some first attestations of *zoeken* 'to search', a verb which does not really denote a mental activity, and one example with *lopen* 'to run', see (242) and (243). Only in the 1950s does the first experience verb *zich vervelen* 'to be bored' show up, see (244).

(242) Ik heb **me suf gezocht**, hijgde hij, waar zat je? (Delphcorp, 1930-1939)
I have myself drowsy searched [...]
'I have been looking everywhere, he panted, where were you?'

(243) Ik hat 't feitelek over die eieren. Ik heb **me** van de week **suf gelopen** om der te kreigen. (Delphcorp, 1930-1939)
[...] I have myself this week drowsy run [...]
'I was actually talking about the eggs. I have been running around like crazy this week to get some.'

(244) Hij behoeft alleen maar te leren strooplikken en nog minder dan anders te laten merken dat hij **zich suf verveelt**. (Delphcorp, 1950-1959)
[...] that he himself drowsy bores
'He only has to learn how to lick someone's boots and how to not let it show as much that he is bored out of his mind.'

From the 1970s onwards, the use of *suf* 'drowsy' starts to approximate the present-day situation that was described in Chapter 4. While several of the verb types are still quite

clearly mental activity verbs and *zich suf piekeren* 'to worry oneself drowsy' in particular has developed into a strong collocation, there are many other kinds of verbs that can be boosted by *suf* 'drowsy', as illustrated by (245) to (247).

(245) Niet door **ons suf** te **sparen** en samen miljarden te vergaren, pak dan wat je kunt! (Delphcorp, 1970-1979)
not by ourselves drowsy to save [...]
'Not by saving every single penny and collecting billions... take what you can get!'

(246) De arbiter **annonceerde zich suf**. Op een gegeven ogenblik zelfs zeventien nullen in successie voor Van Bracht. (Delphcorp, 1970-1979)
the referee announced himself drowsy [...]
'The referee was calling announcement after announcement. At one point, he had to announce seventeen zeroes in a row for Van Bracht.'

(247) Dan zei de Bekkerveldtrainer nog over de Poolse coach: hij heeft **zich suf gewisseld** om die set in handen te houden. (Delphcorp, 1970-1979)
[...] he has himself drowsy changed [...]
'About the Polish coach, the trainer of Bekkerveld said: he kept switching out players in an attempt to keep control of the game.'

The development is summarised in the timeline in Figure 5.18, the star marking the first attestation in Delphcorp.

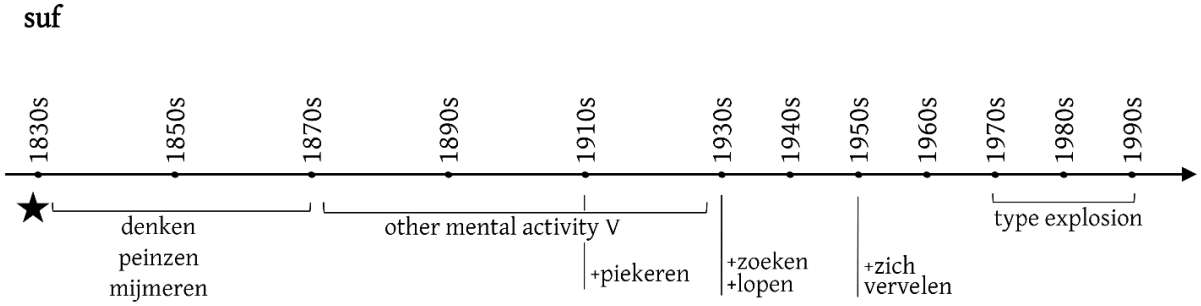


Figure 5.18. Timeline summary of semantic expansion of *suf*

While the diachronic development of *suf* 'drowsy' is an example of far-reaching semantic expansion, we also find evidence of more subtle semantic expansion with some intensifiers that are still semantically constrained in present-day Dutch (cf. Ch4, §4.3.1.2). An interesting example here is *het vuur uit de sloffen* 'the fire out of the slippers', which appears to form an exclusive collocation with *lopen* 'to run' in SoNaR-NL and was therefore argued to impose a lexical constraint on the verb slot in Chapter 4. In Delphcorp as well, all 25 occurrences of *het vuur uit de sloffen* between the 1870s (its first attestation) and the 1950s were with the verb *lopen*. From the 1950s onwards, however, we start seeing other verbs in the verb slot, viz. *draven* 'to trot', *fietsen* 'to cycle', *rijden* 'to ride/drive' and *trappen* 'to pedal'. Although they denote different kinds of activities, all of these activities involve the use of the feet (related to the element of slippers) and are generally also performed with a certain speed (related to the element of fire, think of the popular image of cartoon figures running so fast that their shoes or feet catch fire). This suggests that

the *lexical* constraint (i.e. only *lopen* 'to run') has relaxed into a *semantic* constraint, allowing for some more flexibility in the verb slot (Zeschel 2012: 7).

(248) Wim van Est, die drommels goed wist, dat twee van zijn maats voorop lagen, **trapte zich het vuur uit de sloffen** om het wiel van de sprintende Magni te houden. (Delphcorp, 1950-1959)
[...] pedalled himself the fire out of the slippers [...]
'Wim van Est, who knew damn well that two of his teammates were in the lead pedalled like crazy to stay in the wheel of the sprinting Magni.'

(249) Bakker **reed zich het vuur uit de schoenen** om geen ronde achterstand op te lopen. (Delphcorp, 1950-1959)
bakker rode himself the fire out of the shoes [...]
'Bakker rode his butt off in order to not fall behind a full lap.'

(250) En hij heeft **zich** vooral zondag **het vuur uit zijn rennerssloffen gefietst** om der wille van het tricot, met succes. (Delphcorp, 1950-1959)
and he has himself mainly sunday the fire out of his cycling shoes cycled [...]
'Especially on Sunday, he cycled his butt off to win the jersey, and succeeded.'

As shown by examples (249) and (250), we also start seeing some variation in the prepositional part of the intensifier, often in accordance with the verb it is used with or to provide a better fit with the context. Other variations are *molières* 'lace-ups', *schaatsen* 'skates', *slofjes* 'little slippers', *spikes* 'spikes', *sportschoenen* 'trainers', *voetbalschoenen* 'soccer shoes', and even some non-footwear items like *schenen* 'shins' or *spaken* 'spokes'. By the 1960s, there are even a couple of examples in which the semantic constraints no longer seem to apply.[62] In case of a verb like *sloffen* 'to shuffle' we still have an activity involving the use of one's feet, but the element of speed is clearly absent, making example (251) sound almost like a contradiction in terms. In the examples (252) to (254), with the verbs *werken* 'to work' and especially *vergaderen* 'to meet' and *praten* 'to talk', it appears that the original semantics of the intensifier impose little to no constraints on the verb at all. Of course, it is possible that the restrictions are not so much eroded, but are instead deliberately ignored to create extra effect.

(251) Paul Carlitz uit Chèvremont **slofte zich** gisteren **het vuur uit zijn sloffen** om voor hen nog tijdig onderdak te vinden. (Delphcorp, 1970-1979)
[...] shuffled himself yesterday the fire out of his slippers [...]
'Paul Carlitz from Chèvremont shuffled around yesterday, trying to find them a place to stay in time.'

[62] Full list of other verbs with *het vuur uit de sloffen* 'the fire out of the slippers' (and variants) in Delphcorp: *draven, fietsen, praten, reizen, racen, rennen, rijden, schoppen, sjouwen, sloffen, spelen, trappen, vergaderen, werken* (for English translations, see the translation list at the beginning of the thesis)

(252) Natuurlijk, ideologisch zal hij **zich het vuur uit de sloffen** moeten **praten**. (Delphcorp, 1960-1969)
[...] he himself the fire out of the slippers must talk
'Naturally, ideologically speaking he will have to talk some serious game.'

(253) Of van iemand die het eenvoudigweg mooi werk vindt om **zich het vuur uit de sloffen** te **werken** voor het skütsjefonds. (Delphcorp, 1990-1995)
[...] to himself the fire out of the slippers to work [...]
'Or someone who simply finds it pleasant to work his butt off for the "skütsje" fund.'

(254) Achterhaald is het beeld van de beminnelijke oudere dame, die **zich het vuur uit de sloffen vergadert** over een onderdak voor thuisloze zwerfpoezen. (Delphcorp, 1990-1995)
[...] who herself the fire out of the slippers meets [...]
'The idea of a lovely old lady who meets with a bunch of people, trying to find shelter for stray cats, is outdated.'

In light of this development, summarised in Figure 5.19, it is quite curious that the SoNaR-NL data set only contains examples with *lopen* 'to run'. It actually seems as if the intensifier, after a brief period of more relaxed use, reverted back to its original lexical constraint. This could be related to the recent decrease in the frequency curve that was observed for *het vuur uit de sloffen* from the 1990s onwards, but it is too soon to say. As was mentioned in Chapter 3, §3.2.1.1, the collocation may be lexicalising into a fully fixed expression with a holistic meaning, given the definition in Van Dale of *het vuur uit zijn sloffen lopen* as 'to put in a lot of effort for something or someone'.

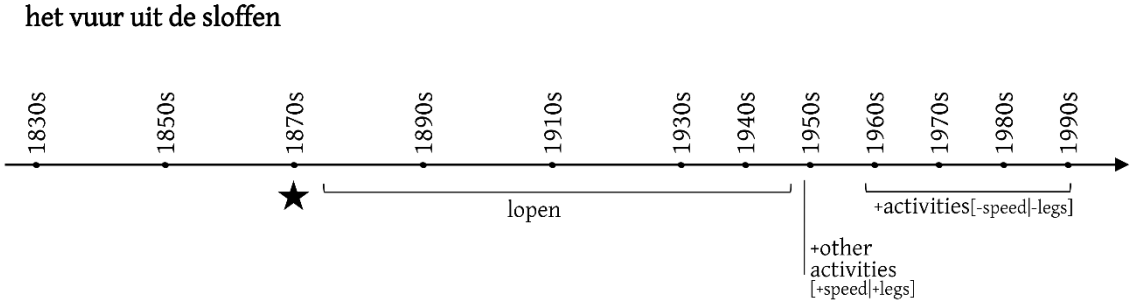


Figure 5.19. Timeline summary of semantic expansion of *het vuur uit de sloffen*

Another intensifier which was found to impose a semantic constraint on its verb slot in Chapter 4 was *de longen uit het lijf* 'the lungs out of the body'. In present-day Dutch, two distinct categories of verbs are combined with this intensifier, viz. verbs of air expulsion or noise emission and verbs of heavy physical activity. What ties these verb classes together is that they both require some physical effort. The first attestation of *de longen uit het lijf* 'the lungs out of the body' in Delphcorp dates back to the 1950s and features the verb *schreeuwen* 'to scream'.

(255) Ook kan niemand u horen, al **schreeuwde** u **zich de longen uit het lijf**. (Delphcorp, 1950-1959)
[...] if screamed u yourself the lungs out of the body
'No one can hear you anyway, even if you screamed from the top of your lungs.'

It appears that the combination with the air expulsion/sound emission verbs was primary.[63] Some earlier exceptions with *lopen* or *rennen* 'to run' notwithstanding (see (256)), the physical activity verbs like *sjouwen* 'to drag' or *fietsen* 'to cycle' are only used with some regularity from the 1980s onwards.[64]

(256) We **lopen ons** met zijn drieën **de longen uit het lijf**. Kreuz en de Jong gaan bij elke aanval mee naar voren. (Delphcorp, 1970-1979)
we run ourselves with us three the lungs out of the body [...]
'The three of us are running our socks off. With each offensive, Kreuz and de Jong also push forward.'

(257) Tonnie Teuben **fietste zich** bijna **de longen uit het pezige lijf**, maar moest te veel werk alleen opknappen. (Delphcorp, 1980-1989)
tonnie teuben cycled himself almost the lungs out of the stringy body [...]
'Tonnie Teuben was cycling his butt off, but he had too much to tackle alone.'

(258) Tijdens de finale van Jumping Amsterdam, **renden** en **sprongen** beide dieren **zich de longen uit het lijf**. (Delphcorp, 1990-1995)
[...] ran and jumped both animals themselves the lungs out of the body
'During the finals of Jumping Amsterdam, both animals ran and jumped as if their lives depended on it.'

Interestingly, the 1990s also contain an example with *lachen* 'to laugh', which is part of the class of emotional verbs in our semantic classification. Of course, *lachen* also involves some kind of sound emission and, in that sense, does not really violate the imposed constraints, but it still is a less prototypical verb than, e.g. *schreeuwen* 'to scream', *zingen* 'to sing' or *juichen* 'to cheer'.

(259) Minutenlang hebben we, zonder eigenlijk een woord te wisselen, in die pikdonkere slaapkamer **ons de longen uit het lijf gelachen**. (Delphcorp, 1990-1995)
[...] ourselves the lungs out of the body laughed
'For minutes we laughed our heads off in the pitchblack bedroom, without saying a word.'

[63] Full list of air expulsion/noise emission verbs with *de longen uit het lijf* 'the lungs out of the body' in Delphcorp: *blaffen, blazen, brullen, gillen, hoesten, juichen, scanderen, schreeuwen, trompetteren, zingen* (for English translations, see the translation list at the beginning of the thesis)

[64] Full list of physical activity verbs with *de longen uit het lijf* 'the lungs out of the body' in Delphcorp: *fietsen, knokken, lopen, rennen, rijden, sjouwen, springen, trainen, trappen* (for English translations, see the translation list at the beginning of the thesis)

The development that *de longen uit de lijf* 'the lungs out of the body' has already undergone in its brief period of existence is summarised in Figure 5.20.

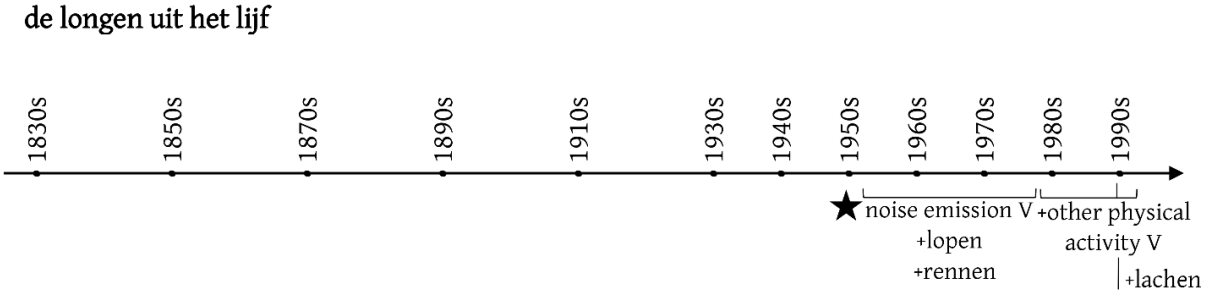


Figure 5.20. Timeline summary of semantic expansion of *de longen uit het lijf*

In present-day Dutch, the intensifier *uit de naad* 'out of the seam' was found to show similar collocational preferences to and share some specific collocates with *de longen uit het lijf* 'the lungs out of the body'. Its early uses in Delphcorp suggest that *uit de naad* 'out of the seam' has also undergone a similar semantic expansion, but in the reverse direction (see Figure 5.21). *Uit de naad* 'out of the seam' was limited to the verb *werken* 'to work' (which is, for that matter, not used with *de longen uit het lijf* 'the lungs out of the body') until the 1970s, at which point it was extended to other physical activity verbs, such as *trappen* 'to pedal' and *fietsen* 'to cycle'.[65] In the 1970s, some noise emission verbs are also added to the list of possible verbs, see (262), but these remain very infrequent, both in terms of tokens and types, compared to the physical activity verbs.[66]

(260) Chrissie **werkte zich uit de naad** voor de lui. Maar of je me nou gelooft of niet, daar deugde ze ook al weer niet. (Delphcorp, 1930-1939)
chrissie worked herself out of the seam [...]
'Chrissie worked very hard for those people. But whether you believe it or not, she wasn't any good at it.'

(261) En die vrouw lacht maar. Daar het ie nou zich voor **uit de naad getrapt**? (Delphcorp, 1940-1949)
[...] himself for out of the seam pedalled
'And the woman kept laughing. Is this why he had pedalled like crazy?'

[65] Full list of physical activity verbs with *uit de naad* 'out of the seam' in Delphcorp: *acteren, dansen, fietsen, knokken, lopen, ploeteren, rennen, rijden, schaatsen, sjouwen, slaan, springen, trappen, vliegen, voetballen, werken* (for English translations, see the translation list at the beginning of the thesis)

[66] Full list of noise emission verbs with *uit de naad* 'out of the seam' in Delphcorp: *blazen, hoesten, schreeuwen, spelen* (for English translations, see the translation list at the beginning of the thesis)

(262) Het is een gezapig zootje dat liever gemakkelijk een paar centjes verdient in een of andere studio, dan **zich** 's avonds ergens **uit de naad te blazen** voor weinig. (Delphcorp, 1970-1979)
[...] than themselves at night somewhere out of the seam to blow [...]
'They are an easygoing bunch that would rather make a couple of easy bucks in some studio, than play their hearts out for little to no money.'

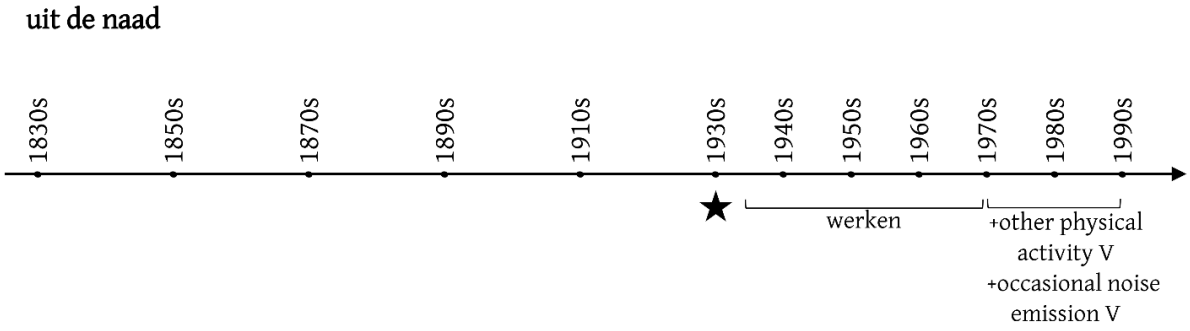


Figure 5.21. Timeline summary of semantic expansion of *uit de naad*

While some of the NP+PP intensifiers (and the PP intensifier *uit de naad* 'out of the seam') have relaxed their collocational constraints, they appear to do so at a much slower pace than the more general adjectival intensifiers that were mentioned earlier. It is quite plausible that the process of semantic bleaching or the development into pure intensifiers is prevented, or at least delayed by their lexical weight or lexical specificity. The same scenario could be assumed for the NP+AP intensifiers. For example, we found that the intensifiers *de handen blauw/stuk/kapot* 'the hands blue/broken' are exclusively combined with *klappen* 'to applaud', an activity which naturally involves the use of one's hands.

(263) Tijdens een emotionele bijeenkomst **klapten** gisteren 1500 stakers **zich de handen stuk** voor deze man. (Delphcorp, 1960-1969)
[...] clapped yesterday 1500 strikers their hands broken
'During an emotional gathering yesterday, 1,500 strikers clapped enthousiastically for that man.'

As the NP+AP intensifiers are overall quite infrequent and most of them do not even appear in every single decennium, however, it is not easy to track their development.

(b) Decrease in collocational range

Thus far, only examples of collocational and semantic *expansion* were considered. It was suggested in Chapter 2 that constructions may also see their collocational range shrink or become confined to particular collocations, which then serve as remnants of their former productivity. This generally happens when a construction is losing ground to another competing construction. Given that the intensifiers are also competing among one another (cf. Ch2, §2.3.2 and §5.3.1 above), we may expect to find that at least some

intensifiers that used to have a wider coverage have lost verb types or have retreated to particular semantic niches. The two examples that will be discussed here allow us to look into such collocational "contraction" in more detail. The intensifier *een aap* 'a monkey' was already shown to be on a declining trajectory in the later decennia in Delphcorp, to the point where it may even have disappeared in present-day Netherlandic Dutch. The intensifier first appeared in the 1910s in combination with the verb *lachen* 'to laugh'. In the 1930s, only 1 out of 8 occurrences were with *lachen*, the other 7 being with the verb *schrikken* 'to be startled'.

(264) Je **lacht je** gewoon **een aap**, meneer; hij kan iedereen nadoen. (Delphcorp, 1910-1919)
you laugh yourself simply a monkey [...]
'You will simply laugh your head off, sir; he can imitate anyone.'

(265) Mopje **schrok zich** toen **een aap**, zette 't op een loopen. (Delphcorp, 1930-1939)
mopje startled herself then a monkey [...]
'Mopje was very startled and started running.'

Together, these verbs continued to dominate the collocational range until the late 20th Century. In the entire data set, *een aap* 'a monkey' was found with only three other verbs besides *lachen* 'to laugh' and *schrikken* 'to be startled', viz. *klappen* 'to clap', *werken* 'to work', and *zoeken* 'to search', see (266) to (268). Strikingly, all three of these were used in the 1990s, at which time *een aap* 'a monkey' was already clearly on the decline with only 8 attestations total (compared to 34 at its peak in the 1950s).

(266) Voetbal-journalist Matty Verkamman had **zich** niet vergeefs **een aap gezocht** naar een oud, gaaf Oranje-shirt met een rood-geel touwtje. (Delphcorp, 1990-1995)
[...] had himself not in vain a monkey searched [...]
'Soccer journalist Matty Verkamman's intense search for an old, intact Orange shirt with a red and yellow ribbon had not been in vain.'

(267) De zaal **klapt zich een aap** om zoveel geld en de kandidaten kijken wat ongemakkelijk rond. (Delphcorp, 1990-1995)
the room clapped itself a monkey [...]
'The room clapped enthusiastically at all the money, and the participants looked around awkwardly.'

(268) Vraag het aan iedere vakman. Willink heeft **zich een aap gewerkt**. (Delphcorp, 1990-1995)
[...] willink has himself a monkey worked
'Ask any professional. Willink worked very hard.'

Rather than gradually contracting its collocational range over time, *een aap* 'a monkey', which was never a very productive intensifier to begin with, appears to have had a brief period of some combinatorial flexibility before disappearing completely in present-day Netherlandic Dutch, as can be seen in the summary in Figure 5.22.

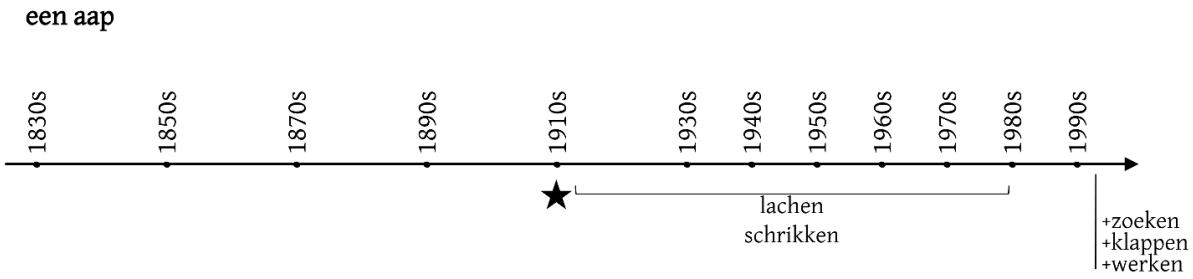


Figure 5.22. Timeline summary of collocational development of *een aap*

The second example illustrates that an intensier may undergo collocational narrowing even if it is not decreasing in overall frequency. In present-day Dutch, the intensifier *wild* 'wild' still has 34 attestations which are almost equally distributed over the two verbs *zich ergeren* 'to be annoyed' and *schrikken* 'to be startled'. The diachronic data show that *wild* 'wild' used to have a wider collocational range. In the 1940s and 1950s *schrikken* 'to be startled' already accounted for 4 out of 10 total occurrences but there was no sign yet of *zich ergeren* 'to be annoyed'. Other verbs were *lezen* 'to read', *lachen* 'to laugh', *gillen* 'to screech', *zich schamen* 'to be embarrassed' and *zoeken* 'to search', see, e.g. (269) to (271).

(269) De raket landde op 300m van de lanceerinrichting in een moeras, zodat het bedieningspersoneel **zich wild** moest **zoeken** om haar terug te vinden. (Delphcorp, 1950-1959)
[...] itself wild must search [...]
'The rocket landed 300m away from the launcher in a swamp, so the operating crew had to search intensely to find it.'

(270) Toen ze de Uitslag gingen omroepen, **gilde** iedereen **zich wild**, toen Hassenforder als winnaar werd bekend gemaakt. (Delphcorp, 1950-1959)
[...] screeched everyone themselves wild [...]
'When they announced the results, everyone screamed their lungs out when Hassenforder was proclaimed the winner.'

(271) Iedereen **lachte zich wild** en niemand durfde ook maar het plan opperen deze "rare dingen" aan te trekken. (Delphcorp, 1950-1959)
everyone laughed themselves wild [...]
'Everyone was laughing hard and no one even dared to suggest actually wearing these silly things.'

In the 1960s, the first occurrence with *zich ergeren* 'to be annoyed' was found, see (272), but it remained relatively infrequent until the 1990s. In addition to the frequent use with *schrikken* 'to be startled', *wild* 'wild' continued to be used with several other verb types from different semantic classes up until the 1980s, as illustrated by (273) and (274).

(272) Ik kan **me** er **wild** om **ergeren** als jonge schilders nu met hun lelijke eend hun contraprestatie gaan halen en dan zeggen dat ze niet kunnen werken omdat ze geen materiaal hebben. (Delphcorp, 1960-1969)
I can myself wild about annoy [...]
'I am so annoyed when these young painters take their ugly deux-chevaux to go pick up their return and then complain about not being able to work because they don't have supplies.'

(273) Je **trapt je wild** en komt bijna niet vooruit. (Delphcorp, 1970-1979)
you pedal yourself wild [...]
'You're pedalling like crazy and you're hardly moving an inch.'

(274) Het Franse staatsgas en elektriciteitsbedrijf **stookt zich** iedere winter **wild** en hoeft bijna niets te betalen. (Delphcorp, 1980-1989)
[...] heats itself every winter wild [...]
'Every winter, the French gas and electricity company can heat as much as it wants, while paying next to nothing.'

Still, it appears that *schrikken* 'to be startled' and *zich ergeren* 'to be annoyed' have gradually ousted all other verbs. By the 1990s, *wild* 'wild' has become confined to those two experience verbs. Although *wild* appears to have decreased in productivity, the fact that it has become part of two conventional fixed expressions, see Figure 5.23, may explain why *wild* increases in overall frequency and it may in fact guarantee its survival (for now).

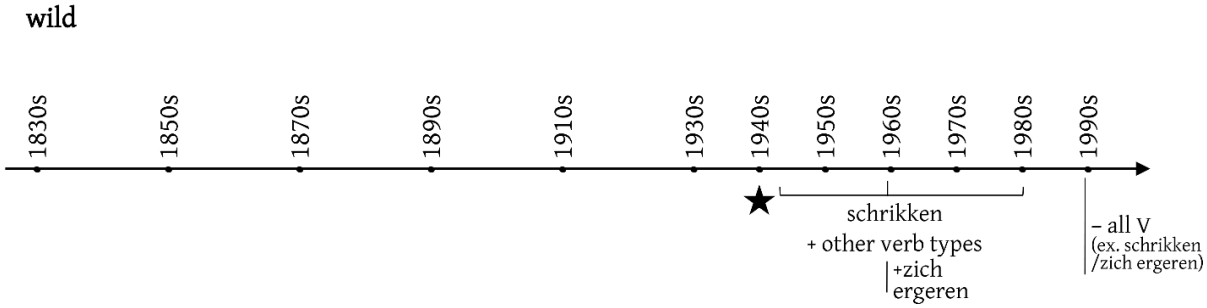


Figure 5.23. Timeline summary of semantic narrowing/conventionalisation of *wild*

To conclude, it appears that the token and type frequency increases that were observed for several intensifiers in §5.3.1 are in many cases parallelled by a steady widening of semantic scope. The examples of a decrease in productivity and/or semantic narrowing are less numerous and, moreover, not always easy to interpret. There is a certain fickleness about the competition between intensifiers that makes it difficult to predict which intensifiers will be affected in what way. The unpredictability of linguistic competition will be further discussed in Chapter 6, §6.2.2. Before turning to what the implications of these findings are for the taxonomic structure of the constructional network, we briefly discuss the issue of semantic coherence from the perspective of the verb.

5.3.2.2 Verbs

The observation from Chapter 4 that verbs are, generally speaking, less likely to impose semantic constraints on the intensifiers they co-occur with, also appears to apply to older stages of Dutch. Still, we will discuss three interesting case studies to show how semantics may be relevant for verbs as well. Several of the intensifiers that were discussed in the previous subsection were first introduced in the construction as part of (near-)exclusive collocations. While the previous paragraph illustrated how the intensifiers have gradually extended to new individual verb types and new semantic classes of verbs, the question can be raised as to what happened to the verbs that were part of these collocations. The intensifier that has undergone the most drastic changes overall is *suf* 'drowsy'. The original collocations *zich suf denken/peinzen* 'to think oneself drowsy' were motivated by mutual semantic compatibility: verbs of mental activity paired up with an adjective that denotes a kind of mental state. Whereas *suf* 'drowsy' has become the most promiscuous intensifier in present-day Dutch, co-occurring with 61 different verb types in SoNaR-NL, the cognitive verbs *denken* and *peinzen* 'to think' have not or barely extended their collocational range to new intensifiers. This clearly shows that, while there are unmistakably important interactions between verbs and intensifiers, individual items are still free to forge their own paths, even if they are (or used to be) part of specific collocations. In a first stage, we see some analogical extensions motivated by semantic similarity in both slots: *suf* 'drowsy' came to be combined with other verbs of mental activity and, occasionally, we also found examples of the mental activity verbs with other "mental state" intensifiers like *het hoofd suf* 'the head drowsy' or *de hersenen suf* 'the brains drowsy', see (275).

(275) Nu veel tyd om **zich** daarover **de hersenen suf** te **peinzen** hadden ze niet want plotseling beleefden zij de schrik van hun leven. (Delphcorp, 1950-1959)
[...] to themselves about it the brains drowsy to ponder [...]
'They didn't have much time to ponder about it because suddenly they got the shock of their lives.'

However, the further expansion of *suf* 'drowsy' beyond the domain of mental activity is not mirrored by the verbs in question: examples with intensifiers that do not denote a mental state in their original semantics are rare. Apparently, the semantic motivation behind the original collocations is still highly relevant in present-day Dutch for the verbs, even though it is no longer so for the semantically bleached intensifier *suf* 'drowsy'. This asymmetry was already touched upon in §5.2, when discussing the ΔP-values of the collexeme analyses: for a collocation like *zich suf piekeren* 'to worry oneself drowsy', the association is heavily asymmetric towards the verb, rather than the intensifier.

The verb *lopen* 'to run', as well, started out in this construction as part of a – semantically speaking – mutually motivated collocation, viz. *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers'. In present-day Dutch, we find that

although *lopen* 'to run' does occur with several intensifiers involving feet or legs (e.g. *de benen uit het lijf* 'the legs out of the body'), it does not impose any clear constraints on its intensifier slot and also freely occurs with all-round intensifiers like *suf* 'drowsy', *rot* 'rotten' and *te pletter* 'to smithereens'. If we now look at the historical development, it seems that *lopen* 'to run' and *het vuur uit de sloffen* 'the fire out of the slippers' heavily relied on each other until the mid-20th Century. With the exception of a single example with *suf* 'drowsy' and *halfdood* 'half-dead' (see (243) above and (276)), *lopen* 'to run' was exclusively used with *het vuur uit de sloffen* 'the fire out of the slippers', and vice versa.

(276) 't Is sport met goede staande honden **zich half dood** te **loopen** in de warme maand September achter de patrijzen. (Delphcorp, 1870-1879)
[...] with good standing dogs themselves half dead to run [...]
'It's a sport to make good dogs run their legs off in the hot month of September to chase partridges.'

We saw that from the 1950s onwards, *het vuur uit de sloffen* 'the fire out of the slippers' ventured beyond the bounds of the exclusive collocation (although it may have reverted back to its original exclusive association in present-day Dutch). At the same time, *lopen* 'to run' started to co-occur with other intensifier types like *lam* 'lame', *te pletter* 'to smithereens', *gek* 'crazy', *te barsten* 'to bursts', etc. By the 1990s, it had become one of the most flexible verbs, combining with 50 different intensifier types from multiple syntactic and semantic categories.[67]

(277) Intussen is het de oprechte snippenjager voor alles te doen om de moeiten van deze jachtwijze, hij **loopt zich lam** door sompe en dras. (Delphcorp, 1950-1959)
[...] he runs himself lame [...]
'In the meantime, the true snipe-hunter will do everything for the sake of the hunt, he constantly runs through swamps.'

(278) Maar ik heb **me** vandaag **te barsten gelopen**, omdat beide partijen de lange pass veelvuldig hanteerden. (Delphcorp, 1960-1969)
but I have myself today to bursts run [...]
'But I have been running around like crazy today, because both teams are constantly giving long passes.'

[67] Full list of intensifier types with *lopen* 'to run' in the 1990s: *blaren, blauw, de adem uit de longen, de benen PREP het gat, de benen PREP het lijf, de benen uit de naad, de blaren, de blaren PREP de voeten, de hakken scheef, de kolere, de longen leeg, de longen uit het lijf, de naad uit het lijf, de pleuris, de poten PREP het lijf, de takken, de zolen PREP de voeten, de zolen uit de sloffen, dood, een hoedje, een ongeluk, een rotje, gek, halfdood, het apezuur, het lazerus, het leplazerus, het licht uit de ogen, het schompes, het schuim op de hiel, het vuur uit de molières/schenen/schoenen/sloffen/slofjes/spikes/ sportschoenen/ voetbalschoenen, in de poeier, in het zweet, kapot, lam, ongelukkig, over de kop, rot, suf, te barsten, te pletter, uit de naad, wezenloos* (for English translations, see the translation list at the beginning of the thesis)

A final interesting case concerns the collocational preferences of *zich ergeren* 'to be annoyed'. Although *zich ergeren* 'to be annoyed' is found with many different intensifier types from multiple syntactic categories and semantic classes in present-day Dutch, it was already briefly mentioned that there are several colour adjectives which appear to be exclusively associated with that particular verb. This association between *zich ergeren* 'to be annoyed' and colours is especially prominent in Delphcorp: approximately 40% of all intensifiers combined with *zich ergeren* 'to be annoyed' in the entire data set are colour-related terms.[68] Although not immediately explicable by the original semantics of the verb or intensifiers, this particular collocational behaviour may be motivated by a symbolic association instead. Given that a lot of the colour terms involve the colour green, it is possible that this colour is associated with the feeling of annoyance (cf. Ch4, §4.1.1.3 on colour symbolism). Still, the fact that we also see other colours, even a rather neutral colour like *grijs* 'grey' in (280), may suggest that the association has extended to other colours besides green as well, at least in this particular construction.

(279) De politie kan alvast beginnen met bekeuringen te maken, aldus schrijft een **zich gras-groen ergerende** lezer. (Delphcorp, 1950-1959)
[...] a himself grass-green annoying reader
'The police can already start giving fines, a very annoyed reader writes.'

(280) Men kan **zich** vaak **grijs ergeren** aan wat men als "vertaling" op het scherm ziet. (Delphcorp, 1960-1969)
one can himself often grey annoy [...]
'One can often be very annoyed by the "translation" that shows up on screen.'

(281) Met name aan Virgil Breetveld **ergerde** de trainer **zich rood en groen**. (Delphcorp, 1990-1995)
[...] annoyed the trainer himself red and green
'The trainer was especially annoyed by Virgil Breetveld.'

With the exception of the conventional intensifiers *blauw* 'blue' and *groen en geel* 'green and yellow', most of these colour adjectives only emerged in the second half of the 20th Century and they are very infrequent. The examples above can therefore be interpreted as analogical extensions or variations of the frequent model intensifiers *blauw* 'blue' and *groen en geel* 'green and yellow'. As was already mentioned in Chapter 4, the journalist can select a specific colour term that is especially fitting in the context in order to create some extra effect. It is probably no coincidence that Virgil Breetveld was playing for the soccer team SVV in the early 1990s, the club colours of which are red and green, see (281).

[68] Full list of colour terms with *zich ergeren* 'to be annoyed' in Delphcorp: *blauw, blauw en groen, geel, geel en groen, grasgroen, grijs, groen, groen en blauw, groen en geel, groen en grijs, paars, paars en groen, rood en groen, zwart* (for English translations, see the translation list at the beginning of the thesis)

The present section has demonstrated that the construction did not only undergo a massive type and token expansion at the highest level of abstraction, there are several individual verbs and intensifiers that have gone through important developments since they were first introduced in the construction. It was argued in Chapter 4 that several of the aspects that were discussed in the productivity section (viz. token frequency, type frequency and semantic range) have an important influence on the hierarchical representation of the constructional network of the intensifying fake reflexive resultative construction. The next section will therefore elucidate how all of these – sometimes substantial, sometimes more subtle - shifts in productivity may have had an impact on the internal structure of the constructional network.

5.4 Reorganisations of the constructional network

Taking together all results presented in the previous sections, we find that the constructional network of the intensifying fake reflexive resultative construction has undergone major transformations since the early 19th Century. The results so far are indicative of a process of schematisation, i.e. the formation of schemas on the basis of concrete instances (cf. §4.4), as well as the further abstraction of the schemas at different levels in the network. Based on the findings in the WNT and the Corpus Literair Nieuwnederlands (Geleyn 2016), we can hypothesise that the network at the turn of the 19th Century consisted of one somewhat productive intensifier-specific subschema [SUBJ V REFL *dood*] – as *dood* 'dead' was the only intensifier that was already found to occur with multiple verbs at the time – and some micro-constructions like [SUBJ *lachen* REFL *stom*] and [SUBJ *lachen* REFL *slap*] (or perhaps a verb-specific subschema [SUBJ *lachen* REFL INT]). In the first decennia in our corpus, we already find some more variation in both the verb and the intensifier slots – which may indicate that some more general subschemas had emerged in the network –, but the data are overall too scarce to form the basis of a meaningful network representation. Even so, it is beyond dispute that the network structure in the (early) 19th Century was much more compact and much less intricate compared to the present-day situation that was presented in Chapter 4. Seeing as how the individual verbs and intensifiers enter into over a thousand different combinations, it stands to reason that we cannot discuss and visualise every single shift that took place within the network over the past two centuries. The main aim of this section is to elucidate that, although the network in general has unmistakably expanded and both verb and intensifier slots have become more schematic, multiple different kinds of shifts may be taking place at lower levels in specific areas of the network. We will therefore primarily focus on a selection of verbs and intensifiers that were shown to have

undergone interesting productivity shifts in the previous section, as well as some more general changes within the construction. In order to capture the most substantial changes within the network, we selected four representative decennia on the basis of the previously established cluster periods, viz. the 1890s to represent the middle of period 1; the 1940s for period 2; the 1970s for period 3 and the 1990s as period 4. That is not to say that no important changes might have happened during other periods, and some idiosyncratic developments or other subtle shifts will necessarily be glossed over. Still, these decennia allow us to capture the most important types of reorganisations within the network. In Chapter 4, we constructed the network in a stepwise fashion, starting out with a very general skeleton and gradually adding further details on the basis of the data that were presented in earlier sections. In this section, we skip the intermediate "building steps" but we do motivate why we have included or excluded certain schemas. These"final" representations of the constructional network (leaving out the lowest level of constructs) will be compared across the selected decennia.

Before we move on to the discussion of some specific cases, we briefly repeat that there is a crucial difference between the way in which the network takes shape in the cognitive-linguistic knowledge of a speaker and the way in which the network is constructed by the linguist (cf. Ch4, §4.4 for discussion). In order to provide a visual representation of the network hierarchy, we (as linguists) are required to make decisions that may affect the eventual hierarchic representation of the network. For example, we need to ask ourselves at which level certain subschemas or generalisations need to be posited, which is no trivial matter. We need to decide whether we will first abstract away from the specific verb or the specific intensifier (that is, do we first get [SUBJ *specified verb* REFL INT] or [SUBJ V REFL *specified intensifier*])? Or which semantic or syntactic generalisations are relevant and at which level do we specify them? However, it was emphasised in Chapter 2 and especially Chapter 4 that these issues are a matter of heuristic choices for the linguist, but they may not be of particular relevance for the language user. While it may not always be possible to reconcile multiple generalisations in one and the same taxonomic representation of the constructional hierarchy, that is not to say that the language user is not able to concurrently capture all these generalisations. In fact, we have argued that he/she has access to multiple different network configurations at the same time. In adopting a multiconfigurational approach to network structures, we may explain certain peculiarities that do not perfectly fit within one representation by referring to interactions with another possible representation. In addition, it may not be practical to *visually* include every piece of information that is associated with specific nodes of the network, but all relevant aspects are of course *part* of the constructional network, if the network is broadly conceptualised as the collection of all knowledge a language user has about the use of this construction. Any representation presented below is a global (and partial) snapshot – based on the corpus data for an entire decennium – of the multidimensional network structure (cf. Chapter 4). For reasons of feasibility, we will

again opt for a general intensifier-centred and a verb-centred representation of the network, but we will discuss other possible configurations when relevant.[69]

5.4.1 1890s

First of all, none of the network representations include semantic constraints at the highest level of schematicity, but this does not mean that there are no relevant restrictions to the verb and intensifier slots at that level. Even in present-day Dutch, not just *any* Dutch verb could be used in the construction: to be considered eligible, the verb needs to have some inherent aspect that can be intensified in one way or another (cf. the absence of stative or unaccusative verbs, cf. Ch4, §4.1.1.1). In order to visually reflect that not all of the categories are already as prominent as they are in present-day Dutch (cf. the semantic expansion discussed in §5.1.1 and §5.1.2), the macro-schema is represented as less schematic or less entrenched (light grey, less thick borders) than it will be in later periods.

Figure 5.24 depicts our proposal for the intensifier-centred network in the 1890s. Given that most of the intensifiers at this point only have one or two attestations, there is insufficient linguistic evidence that speakers have already formed an abstraction over these specific instances at this point – that is, they are all situated at the micro-construction level. As some of these micro-constructions do occur with a certain token frequency, they are assumed to be entrenched to some degree, as indicated by the bold borders. Only the attestations with the intensifier *dood* 'dead' appear to have given rise to an overarching abstraction: its collocates show enough variation to warrant a higher-order subschema in which only the intensifier is lexically specified. The intensifier *suf* 'drowsy' is part of three strong collocations, represented as entrenched micro-constructions at the bottom of the hierarchy. As the three verbs are very similar in meaning – they all denote a mental activity –, it is not impossible that these instances have also triggered a first abstraction in the form of [SUBJ $V_{\text{mental activity}}$ REFL *suf*], but there is no way of knowing whether three verb types is enough for such a generalisation to be formed or not – hence the dashed lines. We included only two higher-order subschemas in which the intensifier is semantically specified. The category of negatively connoted states is more entrenched and more schematic than the inalienable possession schema, not only because it is represented by more intensifier types, but also because it overarches at least one extra intermediate level. That is, while the inalienable possession subschema is immediately formed as a generalisation over micro-constructions, the

[69] Recall from Chapter 4 that the intensifier-centred network is built around intensifier-specific subschemas and higher-order generalisations on the basis of intensifier properties, whereas the verb-centred network is built around verb-specific subschemas and perceived similarities between the verbs that are used in the construction.

negative state subschema has a more complex hierarchic structure which also encompasses the intermediate subschemas [SUBJ V REFL *dood*] and perhaps [SUBJ $V_{mental\ activity}$ REFL *suf*]. Although there is already one micro-construction with an intensifier that will later be considered as a (not very prototypical) member of the diseases category, viz. *een bult* 'a hump', this category is probably not yet recognised by the language user in the absence of other similar cases. Note that we could also have opted for a formal definition of the higher-order subschemas by specifying the syntactic (rather than the semantic) category of the intensifier. It has already been mentioned that the lexical weight of the intensifier, which is related to its syntactic category, appears to have an impact on its collocational freedom, in that NP+PP and NP+AP intensifiers are generally much more limited in their use than AP intensifiers (cf. also Chapter 4). Yet, we have opted for semantically defined subschemas here because we want to show how the *semantic* expansion of the construction over time is reflected in the network structure. At any rate, given that there is a substantial overlap between the semantic and the syntactic categories, in that most of the negatively connoted states are adjectives and the inalienable possession intensifiers are either NP+PP or NP+AP intensifiers, the general lay-out of the network – whether with semantically or syntactically defined schemas – would not differ all that much. Of course, we would like to emphasise once more that it is not a matter of "either-or": the language user has access to multiple possible configurations of the network, depending on the generalities he/she perceives and on his/her linguistic experience.

Figure 5.25 provides a possible verb-centred representation of the constructional network, starting out from the same micro-constructions.[70] We have again opted to focus on semantic generalisations, but it is of course equally possible that language users are also sensitive to the reflexivity of the verb, especially since the group of *zich schamen* 'to be embarrassed', *zich ergeren* 'to be annoyed' and *zich vervelen* 'to be bored' was already somewhat frequent at the time. As all of the inherently reflexive verbs are also emotional verbs, an intermediate subschema specifying the reflexivity could be added between [SUBJ $V_{emotional}$ REFL INT] and [SUBJ *schamen* REFL *de ogen uit het hoofd*], but the reflexivity may also be specified at an entirely different hierarchic level, depending on which kinds of verbal properties are first perceived as relevant by the language user. In the 1890s, *lachen* 'to laugh' is the only emotional experience verb that already occurs with a diverse set of intensifiers, hence the subschema [SUBJ *lachen* REFL INT]. With respect to the cognitive experience verbs, it was already mentioned in §5.3.2 that the attraction between *suf* 'drowsy' and mental activity verbs used to be fully mutual: just like *suf* was exclusively associated with mental activity verbs, the mental activity verbs are only used

[70] Reminder: the order in which the micro-constructions and subschemas are presented from left to right is not meant to be meaningful, it is mainly chosen for practical reasons.

with *suf*. In that regard, the subschema [SUBJ $V_{cognitive}$ REFL INT] may actually be too general. Perhaps we should instead add [SUBJ $V_{mental\ activity}$ REFL *suf*] (cf. the intensifier-centred network in Figure 5.24) in the verb-centred network, but it is not entirely clear what the appropriate level of abstraction for this subschema is in the verb-centred network. Given that the verb slot is not lexically specified, we would expect this subschema to be at a higher level than, e.g., [SUBJ *lachen* REFL INT] but the intensifier is generally only specified at the micro-construction level in the verb-centred network. Within a multiconfigurational or multirepresentational approach to network structures, it is not really essential to force all potentially relevant generalisations into one taxonomic representation: certain kinds of collocational preferences can perhaps be best explained by referring to interactions between multiple configurations of the network (in this case, an interaction between the intensifier-centred and the verb-centred network for the mental activity verbs). In addition to the category of experience verbs, the physical activity verbs are also well-represented in terms of different types (e.g. *knokken* 'to fight', *lopen* 'to run', *ploeteren* 'to plod', *hollen* 'to run', etc.). With the exception of *werken* 'to work', which is quite frequent and displays some combinatorial flexibility, most of these verbs are relatively infrequent and have not given rise to a more abstract subschema yet. The frequent occurrence of the collocation *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' is represented by a very entrenched micro-construction, just like in Figure 5.24. The subschema [SUBJ $V_{noise\ emission}$ REFL INT] is already included because multiple noise emission verbs had been recruited at the time, e.g. *brullen* 'to roar', *juichen* 'to cheer', *schreeuwen* 'to scream', *zingen* 'to sing', etc. but none of them are particularly frequent or flexible, hence the absence of entrenched micro-constructions or intermediate verb-specific subschemas. Several members of the other semantic classes, like the communication verbs, the consumption verbs and more general activity verbs are already attested in the late 19th Century but they are quite likely not diverse enough to have given rise to a higher-level semantic subschema. If we were to include such instances in Figure 5.25, they would not be motivated by a lower-level subschema but instead be directly licensed by the macro-schema [SUBJ V REFL INT] (cf. Ch4, §4.4). We will see whether this has changed about half a century later in the next paragraph.

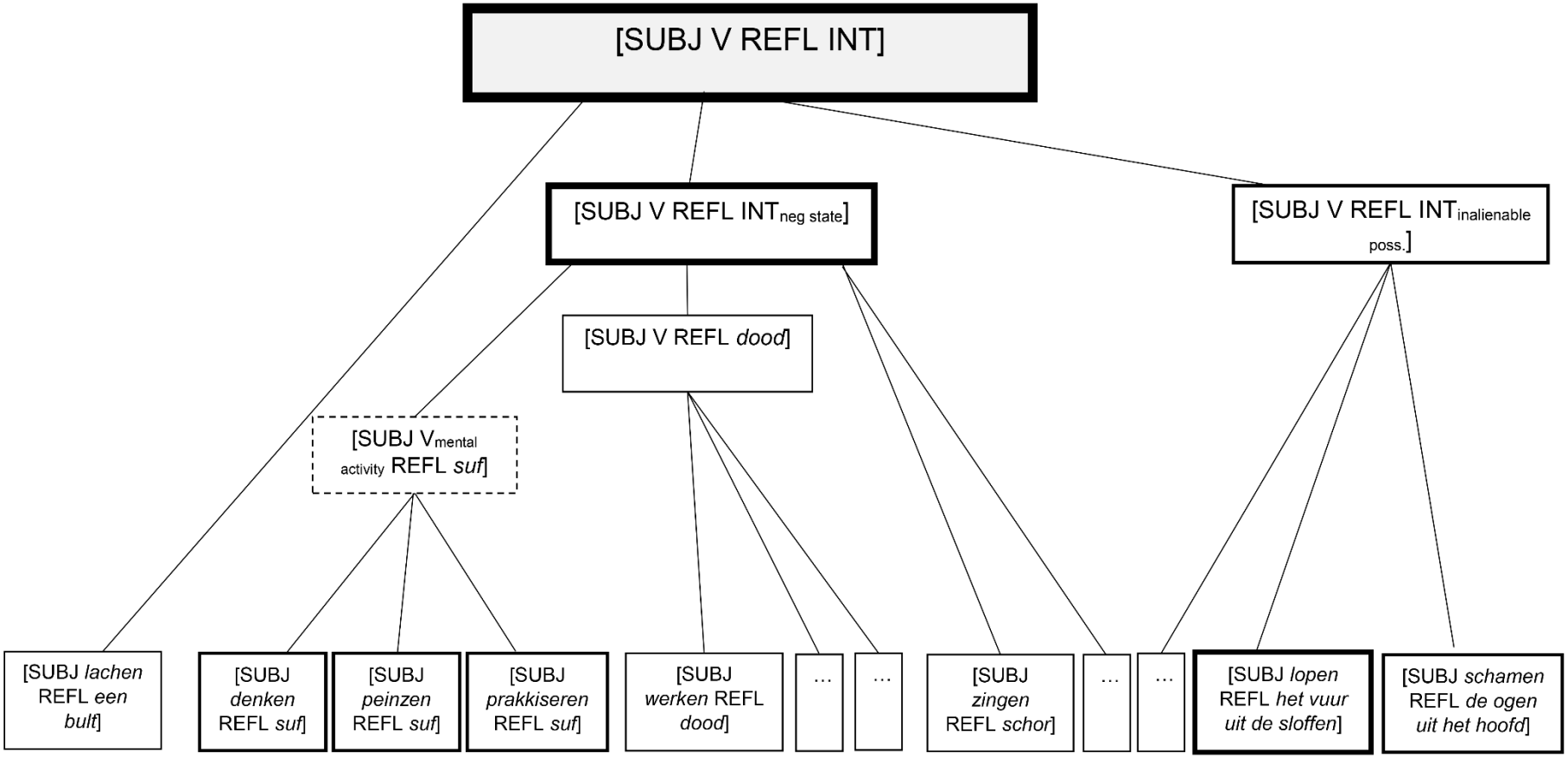


Figure 5.24. Intensifier-centred network representation of the intensifying fake reflexive resultative construction in the 1890s

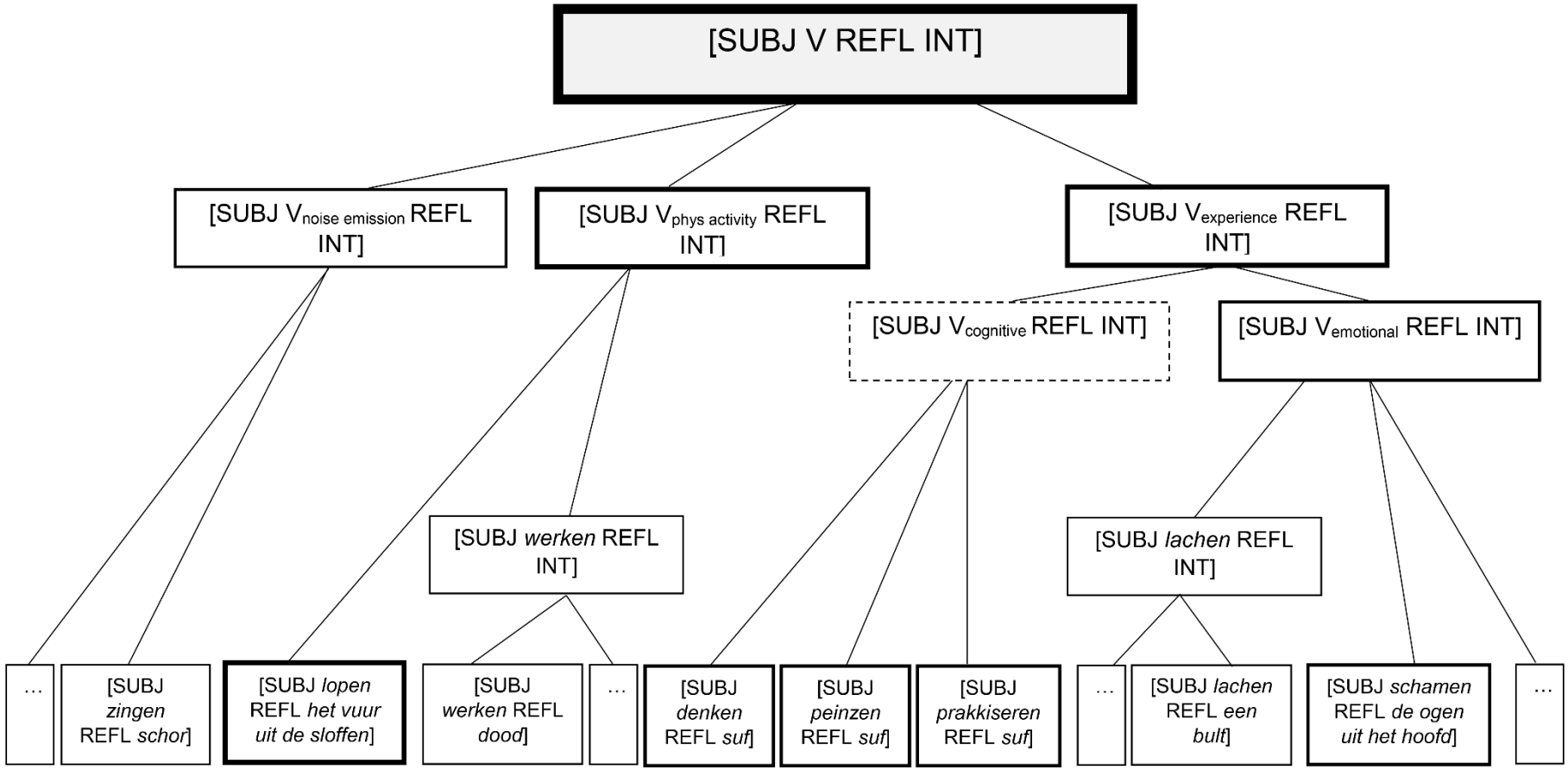


Figure 5.25. Verb-centred network representation of the intensifying fake reflexive resultative construction in the 1890s

5.4.2 1940s

About half a century later, the macro-schema has become somewhat more schematic and entrenched, as the network has become more complex and both the intensifier and verb types have nearly doubled in frequency.[71] In general, there are several parts of the network that have not changed all that much. The intensifier *dood* 'dead' has attracted some new verb types from different semantic categories, but given that the subschema [SUBJ V REFL *dood*] had already been formed in the 1890s, its network representation has not really changed aside from a slight increase in entrenchment. The situation for *het vuur uit de sloffen* 'the fire out of the slippers' has also remained unchanged, as it is still exclusively paired up with the verb *lopen* 'to run' in the 1940s. *Suf* 'drowsy' still has the same three entrenched micro-constructions (visually simplified for reasons of space). As it has also been sporadically occurring with other mental activity verbs (including *piekeren* 'to worry'), there is now sufficient linguistic evidence to assume the existence of a semantically constrained mental activity subschema. Of course, several new intensifiers have been added to the overall repertoire between the 1890s and the 1940s (i.e. new low-level micro-constructions) but most of these are too infrequent to have led to the emergence of more abstract subschemas in Figure 5.26. Still, there are a number of new additions that we need to keep track of, viz. *rot* 'rotten' and *wild* 'wild', which had just been introduced as new (hapax) intensifiers at the time. In addition, we have included two colour intensifiers. The entrenched micro-construction [SUBJ *ergeren* REFL *groen en geel*] represents the strong collocation *zich groen en geel ergeren* 'to annoy oneself green and yellow'. In contrast, *blauw* 'blue' may already be situated at the subschema level (i.e. [SUBJ V REFL *blauw*]) because it is found to co-occur with four semantically diverse verbs, viz. *betalen* 'to pay', *schrijven* 'to write', *schrikken* 'to be startled' and *zich ergeren* 'to be annoyed'. Several other colour terms already occur in the construction, i.e. *geel en groen* 'yellow and green', *groen* 'green', *groen en blauw* 'green and blue', *groen en grijs* 'green and grey', etc. Given their near-exclusive association with *zich ergeren* 'to be annoyed' (with the exception of *blauw* 'blue'), however, it is unclear whether the language user really recognises the colour terms as a productive semantic category of intensifiers at the same level of the inalienable possession intensifiers or the negatively connoted states – for that reason, no subschema [SUBJ V REFL INT_{colour}] is added to the network. The limited combinatorial flexibility of these colour terms could be captured by a subschema [SUBJ *ergeren* REFL INT_{colour}], but this subschema is not easily fit into the current network representation (just like the subschema [SUBJ $V_{mental\ activity}$ REFL *suf*] was not included in

[71] Some parts of the network have been visually simplified to create room for more interesting shifts (e.g. the omission of [SUBJ *schamen* REFL *de ogen uit het hoofd*] and [SUBJ *lachen* REFL *een bult*]).

Figure 5.25). As we will show below, it is possible to account for this collocational preference by positing an interaction with the verb-centred representation of the network, where a subschema like [SUBJ *ergeren* REFL INT_{colour}] is more straightforwardly fit into the hierarchy. In addition to *een bult* 'a hump', some new disease terms have entered the construction (e.g. *de pest* 'the plague', *een stuip* 'a spasm', *een beroerte* 'a stroke' and the first fictitious disease *het apezuur* 'the monkey-acid'). Unlike the colour terms, these diseases can be combined with different verbs so it is possible that the higher-order disease-subschema was already established around the mid-20th Century (albeit not very entrenched yet). Crucially, as most of these disease intensifiers are very (type and token) infrequent, they are mostly situated at the micro-construction level. That is to say, the creation of the subschema [SUBJ V REFL $INT_{disease}$] has probably arisen as a generalisation over multiple micro-constructions with different verbs and different disease types, even in the absence of intermediate subschemas of the type [SUBJ V REFL *specified disease intensifier*].

If we look at the verb-centred network in Figure 5.27, we see several new subschemas compared to the network in the 1890s. In the category of physical activity verbs, the productive subschema with *werken* 'to work' has come to attract more verb types: its position in Figure 5.27 is unchanged but it has become more entrenched. The verb *lopen* 'to run' gradually starts to emancipate itself from the exclusive collocation with *het vuur uit de sloffen* 'the fire out of the slippers'. Several of the emotional experience verbs have also noticeably widened their collocational range to include more diverse intensifier types, thus giving rise to a subschema in which only the verb is lexically specified and the intensifier slot is open. The verb *schrikken* 'to be startled' was only added in the early 20th Century but is already used with several different intensifier types by the mid-20th Century. While Figure 5.27 only includes the reflexive verb *zich ergeren* 'to be annoyed', subschemas have likely also emerged for the other frequent reflexive verbs *zich schamen* 'to be embarrassed' and *zich vervelen* 'to be bored' at this point. What is so interesting about *zich ergeren* 'to be annoyed', is that – while it can co-occur with a diverse set of intensifiers – its distribution contains a clear semantically coherent group of colour intensifiers (cf. supra). This collocational preference is visualised as an intermediate subschema [SUBJ *zich ergeren* REFL INT_{colour}]. In the category of cognitive experience verbs, we now find several cognitive verbs occurring with other intensifiers than *suf* 'drowsy' (e.g. *zich ziek redeneren* 'to reason oneself sick'), which is why we have replaced the dashed lines around the $V_{cognitive}$ subschema by a full line. The category of noise emission verbs continues to attract new members, but all of them are infrequent and their use is mostly limited to one or two intensifiers. Of course, other infrequent verb types, including some sporadic communication verbs, continue to join the distribution at the micro-construction level as well. As they are not (yet) motivated by any intermediate subschemas in the verb-centred network, they would be immediately licensed by the [SUBJ V REFL INT] pattern at the highest level of abstraction.

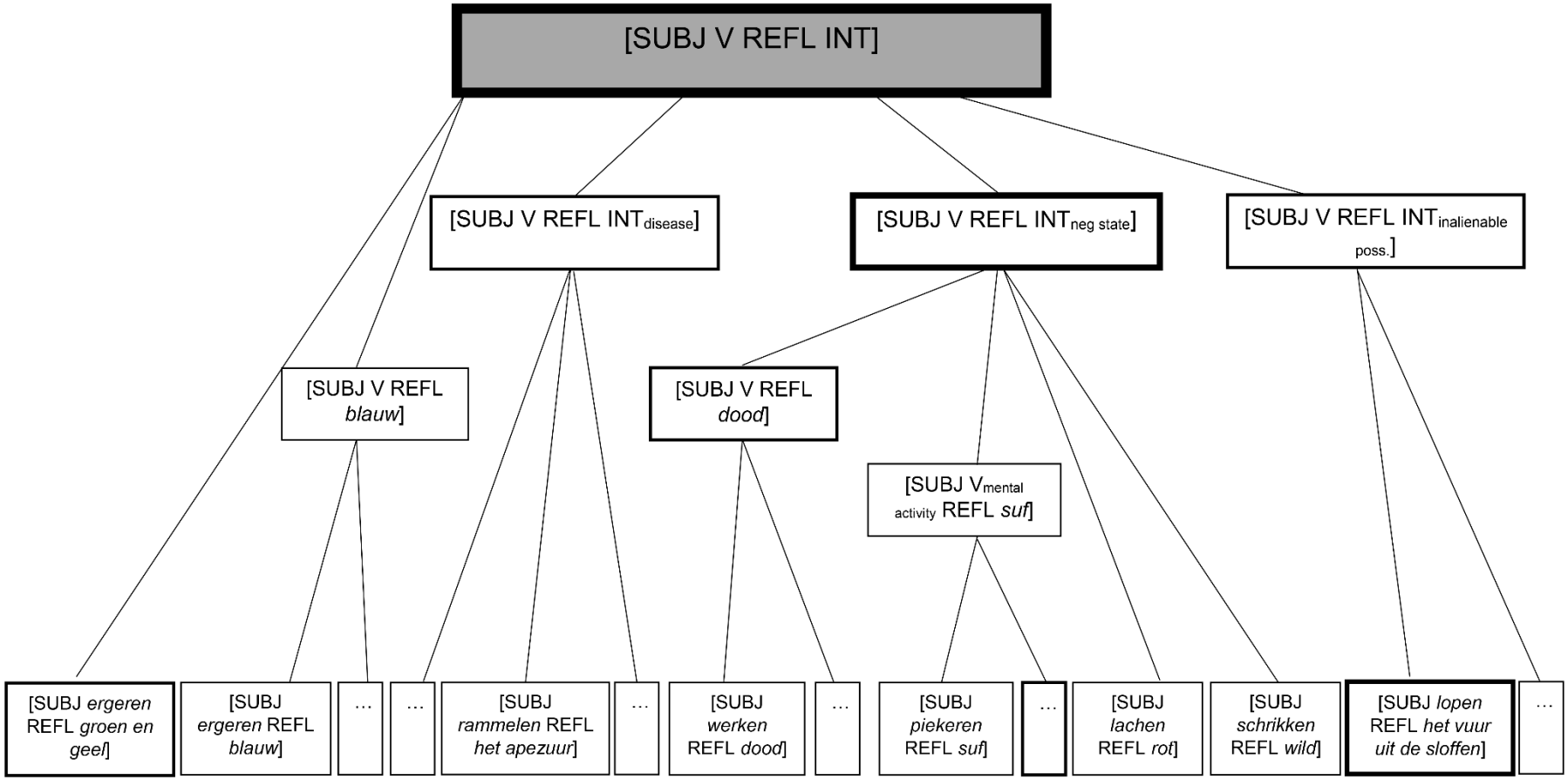


Figure 5.26. Intensifier-centred network representation of the intensifying fake reflexive resultative construction in the 1940s

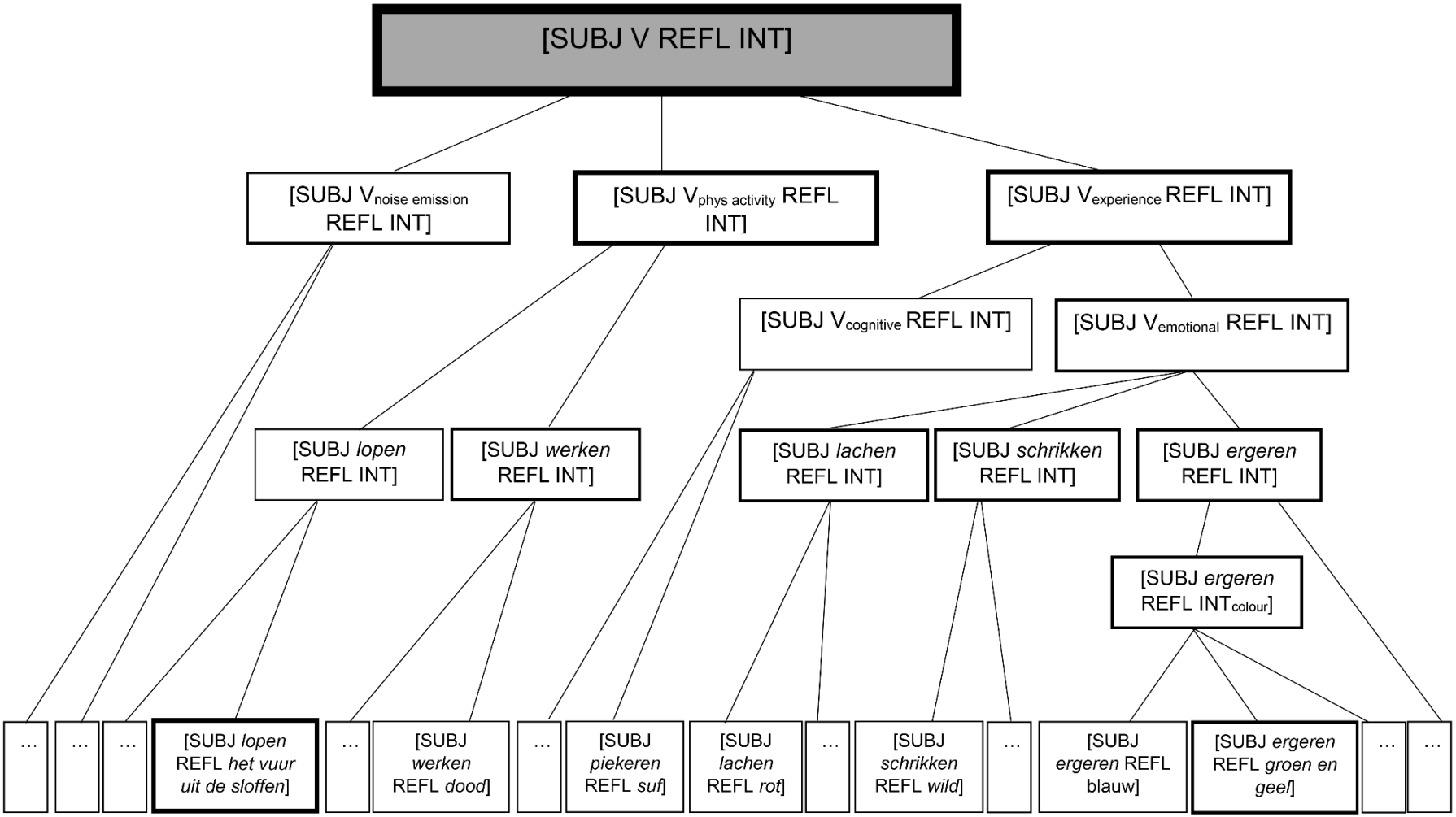


Figure 5.27. Verb-centred network representation of the intensifying fake reflexive resultative construction in the 1940s

5.4.3 1970s

By the 1970s, the macro-schema has again slightly increased its degree of schematicity, as the network continues to expand and the internal structure of the network becomes increasingly complex. The intensifier-centred network in Figure 5.28 has certainly come to look more elaborate than in Figure 5.24 and Figure 5.26. This is especially obvious if we take [SUBJ V REFL *dood*] as our reference point. In §5.3.1.1 *dood* 'dead' was observed to have undergone relatively little change in its position in the global productivity graph, compared to some of the other frequent intensifiers. In the network representation, it is virtually the only schema in the network that has remained largely unchanged since the late 19th Century. To the left of *dood* 'dead', there are some changes in the degree of entrenchment at different levels of abstraction. At the micro-construction level, the combination of *blauw* 'blue' and *zich ergeren* 'to be annoyed' has developed into an entrenched collocation. The abstract semantic schema [SUBJ V REFL INT$_{disease}$] has also become more entrenched as many new disease intensifiers have been added to the repertoire. In the rightmost area of the network, we see that *het vuur uit de sloffen* has come to welcome some new verbs to its collocational range, e.g. *racen* 'to race' or *trappen* 'to pedal'. In §5.3.2 it was argued that *het vuur uit de sloffen* 'the fire out of the slippers' gradually starts to relax its lexical constraint and instead poses a semantic constraint [+legs/+speed] on its verb slot. Some exceptions (such as *sloffen* 'to shuffle', which is [–speed]) are of course not excluded, but they may be deliberate overrides of the semantic constraints, licensed by the most abstract schema. Due to lack of space, we did not include any of the variations on the footwear in the PP part. We suggested in Chapter 4 that such variations can be represented in the network through a direct horizontal link with the model they are based on. Of course, they are also licensed by some higher-level schema in the network, but the horizontal link is meant to emphasise that these micro-constructions have *entered* the network as low-level analogical extensions, rather than as direct instantiations of an intermediate subschema. Concretely, a micro-construction like [SUBJ *lopen* REFL *het vuur uit de voetbalschoenen*] 'the fire out of the soccer shoes' should be primarily regarded not as a new instance of [SUBJ V REFL INT$_{inalienable\ possession}$] (or even of [SUBJ *lopen* REFL INT] in the verb-centred network in Figure 5.29, infra), but as a very local extension on the basis of the entrenched micro-construction [SUBJ *lopen* REFL *het vuur uit de sloffen*]. Over time, such analogical extensions may of course give rise to an overarching schema (cf. the concepts of item-based generalisations and incipient productivity in Zeschel 2012). It is not inconceivable that the different types of footwear at some point become subsumed by an intermediate schema like [SUBJ V REFL *het vuur uit de* N$_{footwear}$]. It is, in fact, quite plausible that this is exactly what happened for the colour adjectives that co-occur with *zich ergeren* 'to be annoyed': originally local variations on *groen en geel* 'green and yellow' (or *blauw* 'blue'), the colour terms eventually became so numerous that they gave rise to the subschema [SUBJ *ergeren* REFL INT$_{colour}$] (cf. supra). A new intensifier in

the network is *de longen uit het lijf* 'the lungs out of the body'. Unlike *het vuur uit de sloffen* 'the fire out of the slippers', it does not immediately form any strong associations with one particular verb, but it is distributed quite evenly across several verbs of noise emission. The most noticeable changes, finally, are found in the middle of the network. Several new intensifier-specific subschemas are established for intensifiers that were introduced in the 1940s. *Rot* 'rotten', for instance, which only had one attestation in the 1940s, has suddenly become one of the most frequent and most flexible intensifiers in the construction. The same can be said for *wild* 'wild': in addition to the micro-construction with *schrikken* 'to be startled' having become more frequent, it can also be used with a variety of other intensifiers in the 1970s. Most strikingly, the intensifier *suf* 'drowsy', which had a long history of being used with just a small set of mental activity verbs, has suddenly expanded its use to a considerable extent. By the 1940s, it had already come to attract several other mental activity verbs, one of which, viz. *piekeren* 'to worry', has now developed into an entrenched micro-construction. Although *suf* 'drowsy' still displays a collocational *preference* for mental activity verbs in the 1970s, it is no longer *restricted* to those verbs: as *suf* 'drowsy' sheds its original semantics and develops into an all-round intensifier that can be paired up with a wide variety of different verb types, a new subschema [SUBJ V REFL *suf*] emerges in the network.

In contrast, the overall organisation of the verb-centred network in Figure 5.29 has not changed all that much since the 1940s. It was already mentioned in §5.1.1 that there is remarkable continuity with respect to the top verbs in the construction. It appears that the main changes within the verb-centred network need to be situated at the lower levels of the hierarchy, which is where new infrequent verbs join the network as low-entrenched micro-constructions. Most of the frequent verbs (some of which are included in Figure 5.27 and Figure 5.29) had already established themselves as abstract, verb-specific subschemas by the 1940s. In the following decades, they continue to grow more frequent and to expand their collocational range to new intensifier types, thus strengthening their representation in the network. Especially *lopen* 'to run', initially confined to the (still frequent) collocation *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers', has come to display remarkable combinatorial flexibility. In addition, some specific micro-constructions have become strong, conventional collocations and are therefore represented as more entrenched. At the higher levels, we added the subschema [SUBJ $V_{communication}$ REFL INT] because the communication verbs have now become sufficiently type frequent to be recognised as a semantically coherent group. The rest category of general activity verbs has also expanded considerably, but given their semantic diversity, they are most likely not subsumed by a semantically specified subschema. Instead, we propose that they are situated at the micro-construction level and are immediately licensed by the most abstract schema.

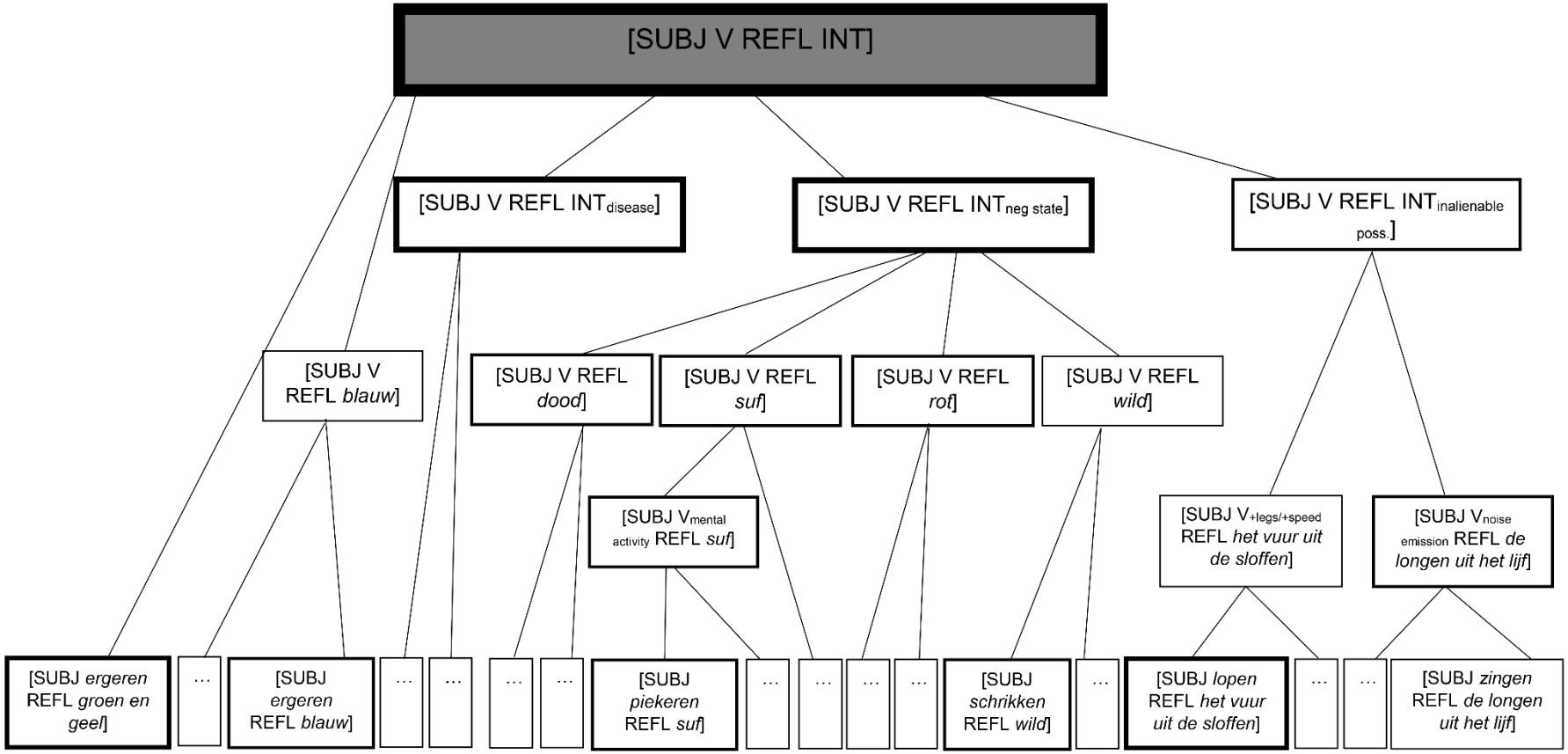


Figure 5.28. Intensifier-centred network representation of the intensifying fake reflexive resultative construction in the 1970s

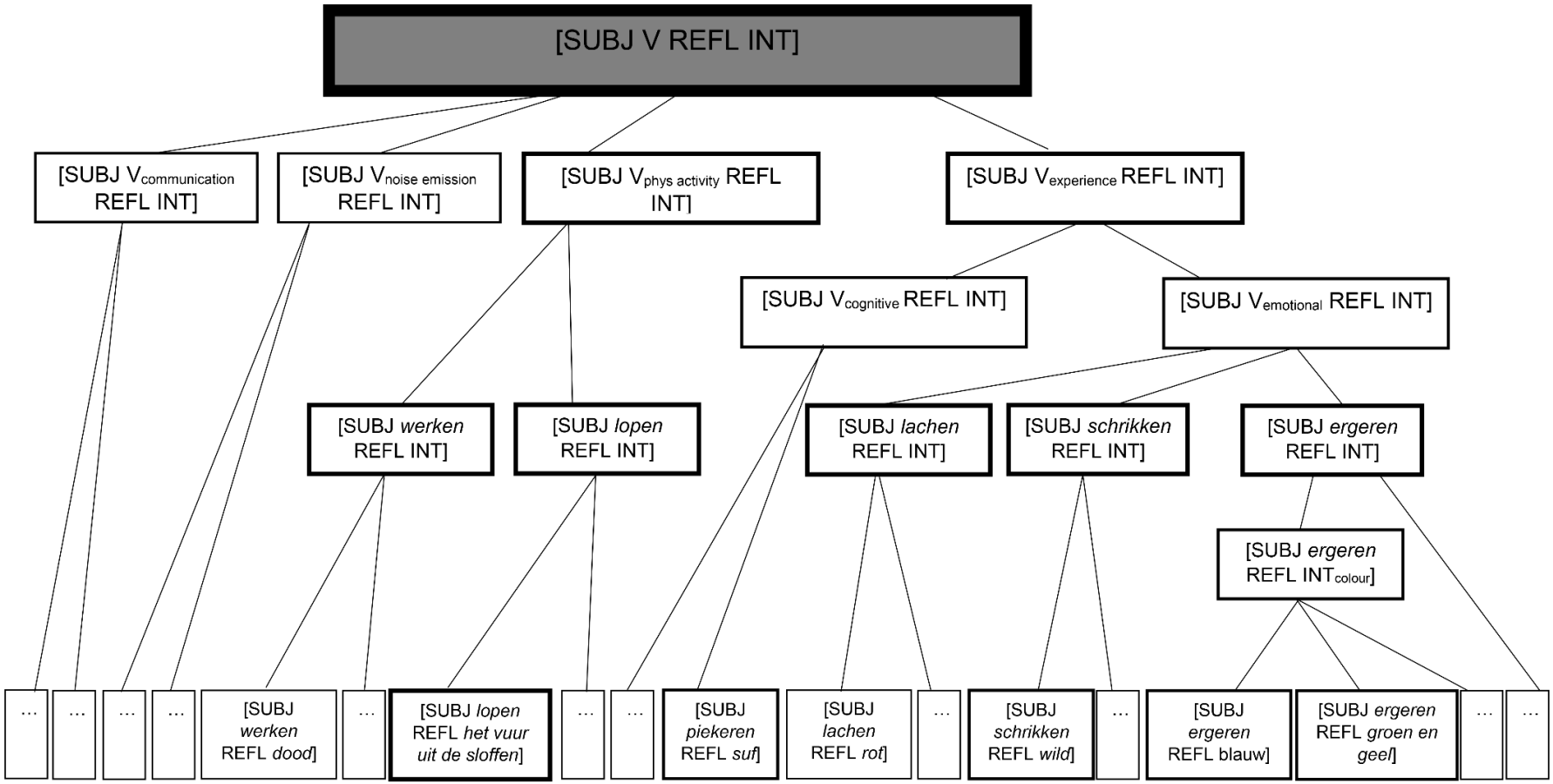


Figure 5.29. Verb-centred network representation of the intensifying fake reflexive resultative construction in the 1970s

5.4.4 1990s

Although several new intensifiers and verbs were added between the 1970s and the 1990s, most of these are very local analogies or subsumed by lower-level schemas in at least one of the network representations. As will be discussed in §5.4.5 and Chapter 6, §6.2.3, these do not necessarily contribute much to the overall schematicity at the highest level of abstraction. That is, the macro-construction has become only slightly more schematic and entrenched compared to the 1970s. As most of the changes between the 1970s and the 1990s are situated at the level of the micro-construction, with new infrequent combinations being added to (or occasionally disappearing from) the bottom area of the network, it was impossible to visually represent these low-level changes in our network representations.

The most important change in the intensifier-centred network in Figure 5.30 concerns the intensifier *wild* 'wild': in the 1990s, *wild* 'wild' has become exclusively associated with *schrikken* 'to be startled' and *zich ergeren* 'to be annoyed' (just like in present-day Dutch). It appears that after a brief period of productivity, in which it could be used with a wider variety of verbs, *wild* 'wild' has retreated to two conventionalised or fossilised collocations. Although the conventionalisation of specific verb-intensifier combinations does not automatically result in the disappearance of the subschema (or vice versa, the emergence of a subschema does not "dissolve" the conventional combination, see Ch6, §6.2.3 for discussion), in this particular case the conventional collocations do appear to have ousted all other combinations, so that we can no longer assume a productive subschema [SUBJ V REFL *wild*]. A very similar fate may be in store for the intensifier *blauw* 'blue', as well. Although it is still combined with 7 verb types in the 1990s, the two collocations *zich blauw ergeren* 'to annoy oneself blue' and *zich blauw betalen* 'to pay oneself blue' account for 46 of the 51 tokens.[72] In Chapter 4, we hypothesised that the fact that *blauw* 'blue' can occur with other verb types aside from *zich ergeren* 'to be annoyed' and *betalen* 'to pay' could be interpreted as *blauw* emancipating itself from these collocations and extending to new verb types. Under that view, it was argued to be curious that the other verb types were not semantically related to the highly frequent models, given that new coinages are generally semantically similar to the already attested types (cf. Suttle & Goldberg 2011, as evidenced by the expansion of *het vuur uit de sloffen* 'the fire out of the slippers' and the early expansion of *suf* 'drowsy'). However, now that we know that *blauw* 'blue' was once a (considerably) more productive intensifier and that the subschema may be weakening, the lack of semantic coherence among the other types makes more sense:

[72] The other verbs are *klikken* 'to click', *lachen* 'to laugh', *oefenen* 'to practise', *schrijven* 'to write' and *solliciteren* 'to apply for jobs'

as relics of a formerly productive schema, they do not necessarily have to display any coherence. In contrast, the subschema [SUBJ V REFL *suf*] has become even more entrenched, as *suf* 'drowsy' is found to occur with 53 different verb types in the 1990s. The mental activity verbs, although still present, are now in the minority. Still, given that they form a rather coherent set compared to the diversity of other verbs, it is quite possible that they are still perceived as a group, although the mental activity subschema may be less entrenched. We also observe that the subschema with *de longen uit het lijf* 'the lungs out of the body' becomes more abstract and schematic, as it appears to have extended its use to a new semantic category of verbs. Although physical activity verbs and sound emission verbs at first blush denote rather different types of activities, they both require a large lung capacity. We captured this similarity in terms of a subschema [SUBJ $V_{\text{lung capacity}}$ REFL *de longen uit het lijf*], although it is possible that this specific constraint may be too ad hoc (cf. Ch4, §4.4.2) or that the shared meaning aspect is not pertinent for the language user.

Our representation of the verb-centred network in Figure 5.31 is mostly the same as in the 1970s, some minor changes in entrenchment notwithstanding. The stability of the network is in line with the results in §5.3.1, which showed hardly any differences in productivity between periods 3 and 4 for most of the frequent verbs. Much like was observed for the intensifier slot above, the changes are primarily situated at low levels in the network and they do not affect our representation here.

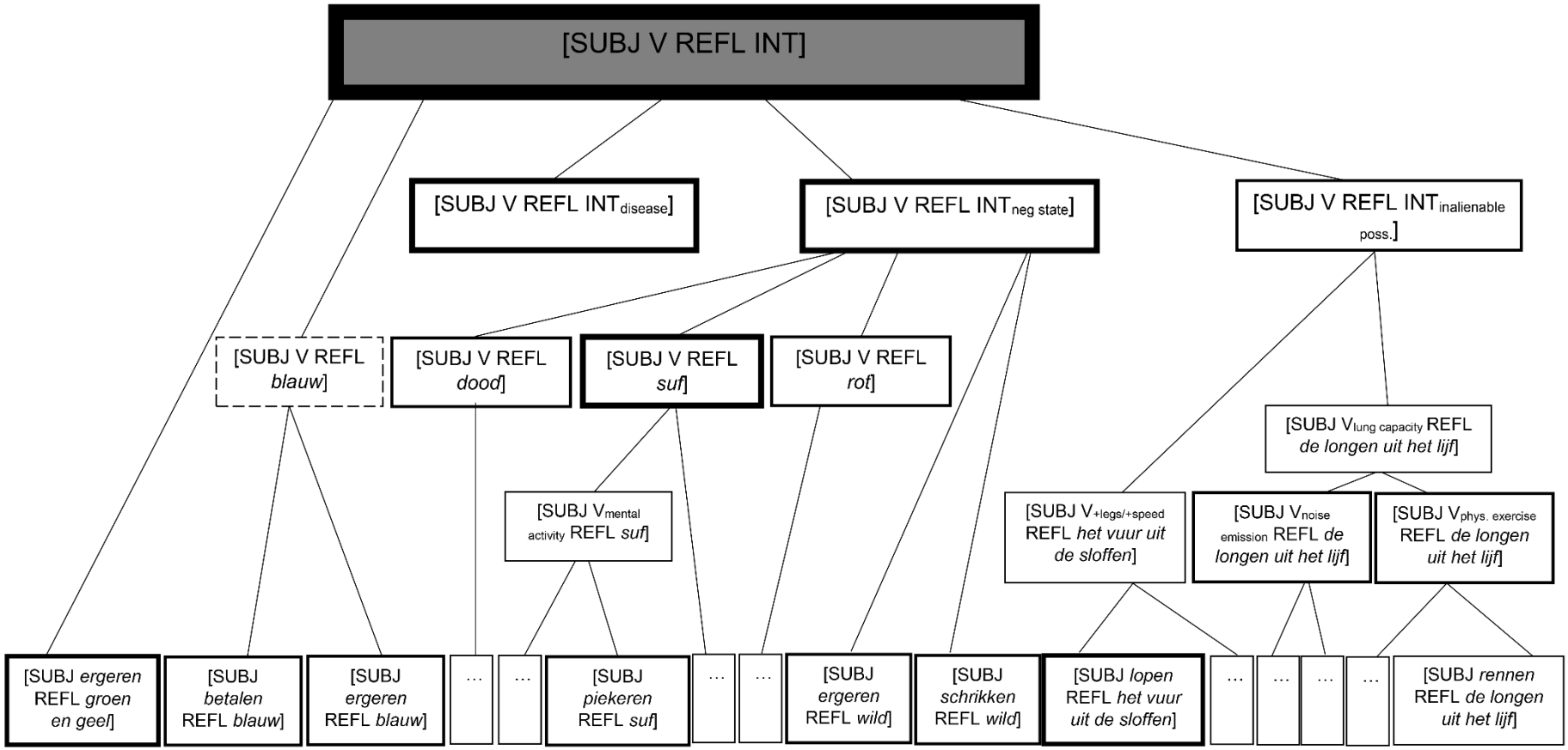


Figure 5.30. Intensifier-centred network representation of the intensifying fake reflexive resultative construction in the 1990s

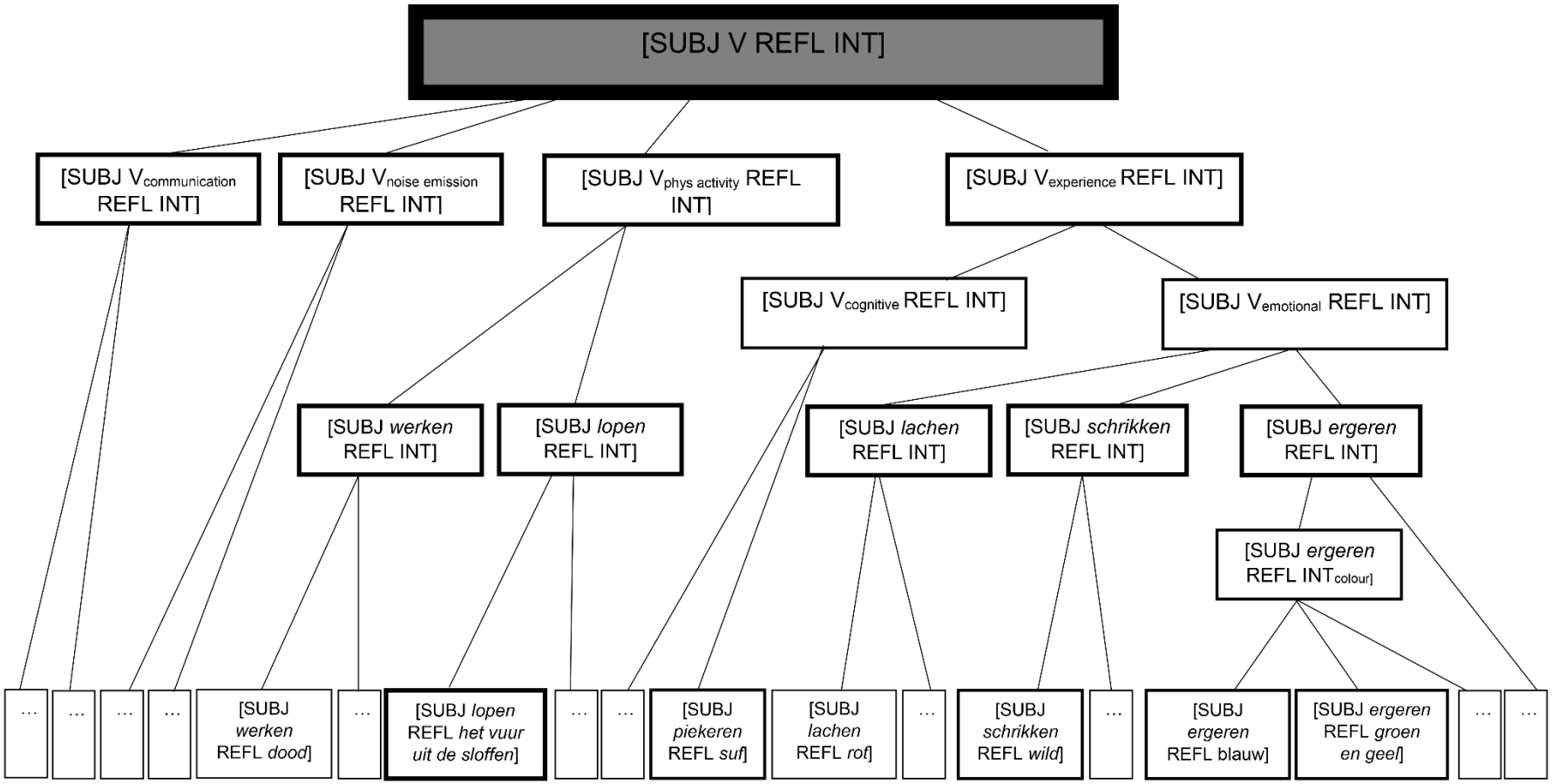


Figure 5.31. Verb-centred network representation of the intensifying fake reflexive resultative construction in the 1990s

5.4.5 Schematisation and conventionalisation: entrenchment at different levels of abstraction

Given the different kinds of shifts that are observed in the constructional network of the intensifying fake reflexive resultative construction, it would be a vast simplification to summarise the historical development of the network merely as expansion. First of all, we saw some typical examples of schema-formation or schematisation, in which a subschema is created as an abstraction over specific instances. Once the subschema has been established, it continues to attract new types, thus increasing its degree of productivity and schematicity and further strengthening its representation in the network. Over time, other generalisations may be formed over these subschemas, giving rise to subschemas at increasingly higher levels of abstraction. At the same time, we also find that specific verb-intensifier combinations at the micro-construction level sometimes increase their frequency and develop into conventional collocations. In linguistic terms, we can claim that both schematisation and conventionalisation are important mechanisms in the reorganisation of the taxonomic representation of the constructional network. Although we have emphasised that it is important to differentiate between the cognitive organisation and a linguist's (visualised) taxonomic representation of the constructional network, we attempted to give a tentative indication of the (cognitive) representation strength of the nodes we included. Where an increase in type frequency was said to consolidate the schematic representation of a subschema, an increase in token frequency would strengthen the cognitive entrenchment of that micro-construction. In Chapter 4, we already showed that the traditional view of Clausner & Croft (1997) on productivity and entrenchment – in which, broadly speaking, either the schema is more entrenched than the specific instances (productivity) or vice versa (non-productivity) – lacks nuance. In the network representation in present-day Dutch, we found that an individual verb or intensifier may be part of a conventional collocation, as well as give rise to an overarching, more abstract subschema. In the previous paragraphs we illustrated that the combination of both (entrenched) productive subschemas and (strongly entrenched) conventional micro-constructions is the natural result of the diachronic development of the construction. We saw that some specific verbs and intensifiers were introduced in the construction as part of a "fixed expression" (i.e. an entrenched micro-construction). Over time, one or both of the elements may emancipate themselves from the collocation and further develop its/their combinatorial possibilities. As a result, a more abstract subschema may emerge but this does not necessarily diminish or nullify the entrenched status of the micro-construction. On the contrary, the conventional collocation may continue to exist and even increase in frequency, hereby becoming even more entrenched (cf. Ch6, §6.2.3).

An interesting recent approach that ties together the "linguistic" processes of schematisation and conventionalisation and the cognitive representation at different levels of abstraction comes from Hilpert (2015a), who suggests a distinction between upward strengthening and cognitive entrenchment. Cognitive strengthening or entrenchment is defined as the entrenchment of a specific linguistic unit after repeated experience of that unit. Upward strengthening, then, is the process by which the experience of a linguistic unit also strengthens the representation of a more (or even the most) abstract schema higher up in the network. Importantly, these processes are of course not mutually exclusive and multiple levels may be strengthened at the same time. It is not always easy to know which level in the network will be strengthened when encountering certain specific instances of the construction. Hilpert (2015a: 137-140) mentions several reasons why some specific instances of a construction may fail to strengthen the abstract, schematic construction. One of the most relevant reasons in light of the current investigation is that the presence of a prominent lower-level subschema may prevent further upward strengthening to higher-order levels. Concretely, a new attestation of, e.g., the intensifiers *suf/dood/rot/...* 'drowsy/dead/rotten' with a new verb will primarily strengthen the subschemas [SUBJ V REFL *suf/dood/rot*], rather than the schema [SUBJ V REFL INT]. Similarly, a new use of the verbs *schrikken/lachen/werken* 'to be startled/to laugh/to work' with a previously unattested intensifier will first strengthen the intermediate subschemas [SUBJ *schrikken/lachen/werken* REFL INT]. A similar claim is found in Perek (2016a: 19), although in talking about schematicity rather than entrenchment, he does not explicitly make any statements about the cognitive status: "an increase in schematicity is not necessary to account for changes in the distribution: new members could be accounted for by being subsumed under low-level schemas, or by analogy with existing exemplars".

A second pertinent factor is the text frequency of the experienced linguistic unit: a new attestation of an already frequently attested chunk is taken to further strengthen the mental representation of that very chunk but not necessarily any higher-order subschemas. This would mean that, for instance, a new occurrence of *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' is expected to first of all contribute to the cognitive strengthening of the already entrenched micro-construction [SUBJ *lopen* REFL *het vuur uit de sloffen*]. While it is not excluded that this strengthening may also affect the subschemas [SUBJ *lopen* REFL INT] or [SUBJ V REFL *het vuur uit de sloffen*], the upward-strengthening-hypothesis assumes that it is unlikely to reach the more abstract levels. This may also explain why subschemas, like [SUBJ V REFL *wild*], can disappear from the network: if one specific instance of a more abstract schema is experienced so frequently, it may no longer strengthen the more abstract schema. Over time, the subschema is weakened or eventually disappears entirely, leaving only entrenched micro-constructions behind. It is argued that upward strengthening towards the highest level of abstraction is most strongly stimulated by "marginal" or truly creative instances of a

construction. In the current investigation, this means that the ideal candidates for upward strengthening are combinations which are directly licensed by the most abstract schema. In the entire data set, there are almost 700 different verb-intensifier combinations that occur only once. As most of these contain either a verb or an intensifier that is quite frequent and flexible, they are more likely to strengthen the subschema in which that particular verb or intensifier is specified rather than the most abstract schema. Still, we do find a couple of examples that may trigger upward strengthening up to the highest level of abstraction. Consider the example of *zich een hoedje sparen* 'to save oneself a hat'.

(282) Hoeveel jongeren **sparen zich** niet **een hoedje** om hun ideaal te verwezenlijken: de aanschaf van een blinkende bromfiets. (Delphcorp, 1980-1989)
how many adolescents save themselves not a little hat [...]
'How many adolescents are not saving all of their money to realise their dream: buying a shiny moped.'

While *een hoedje* 'a little hat' is quite token frequent, the near-exclusivity with *schrikken* 'to be startled' suggests that there is no superordinate subschema [SUBJ V REFL *een hoedje*] and *sparen* 'to save' is also too infrequent to give rise to any intermediate subschemas. Given the lack of clear semantic or formal similarity between *schrikken* 'to be startled' and *sparen* 'to save', we argue that this combination is not a typical analogical extension (unlike, e.g., *zich een hoedje verschieten* 'to startle oneself a little hat', where *verschieten* is a synonym of *schrikken*). Instead, it could be interpreted as an on-the-fly combination of a random verb and intensifier, or, in other words, as a direct instantiation of the most abstract pattern [SUBJ V REFL INT]. Even better examples are those combinations in which both the verb and the intensifier are (relatively) infrequent, although these are admittedly quite rare, see (283) and (284).

(283) Meun zegt niet benauwd te zijn voor de intensieve AID-controles, ook al **controleren** ze **zich de rambam**. (Delphcorp, 1990-1995)
[...] even if control they themselves the rambam
'Meun says that he is not afraid of the intensive AID-controls, even if they check everything extensively.'

(284) Toen **rekten** honderdduizend toeschouwers **zich de nekspieren uit het lid**, om de ballons [...] zo lang mogelijk te volgen. (Delphcorp, 1940-1949)
then stretched one-hundred-thousand spectators themselves the neck-muscles out of the joint [...]
'One hundred thousand spectators were stretching their heads to follow the balloons for as long as possible.'

It was observed in Chapter 4 that language users are sometimes (intentionally or not) "unconventional with conventional means", in which case they combine two token frequent verbs and intensifiers, even though their specific intercombination is not really conventional. However, they do not often seem to be "unconventional with

unconventional means": although the construction contains a considerable proportion of both hapax verbs and hapax intensifiers, these low-frequent items do not always readily combine amongst one another. Language users are much more likely to use (or introduce) infrequent (new) intensifiers with frequent verbs, or mutatis mutandis, infrequent verbs with frequent intensifiers, rather than pairing up two infrequent items. From a diachronic perspective, then, the recent expansion of the intensifying fake reflexive resultative construction is mainly explained by a number of highly frequent verbs and intensifiers extending their use to infrequent, previously unattested items. This brings us back to our earlier hypothesis that the expansion in the most recent decennia primarily strengthens the verb-specific and intensifier-specific subschemas but, according to the upward-strengthening-hypothesis, barely reaches the more abstract construction.

It is important to highlight that the upward-strengthening-hypothesis makes certain cognitive assumptions that need to be put to the test. Concretely, the fact that a specific verb-intensifier combination occurs significantly more often than would be expected on the basis of their individual frequencies suggests that it has acquired the status of a "conventional combination", which is quite likely to be entrenched as such in the cognitive representation of the network as well. While it seems reasonable that a new attestation of this specific combination further strengthens the representation of that combination, the assumption that it does not – at the same time – strengthen some higher-order subschema is less uncontroversial. On the basis of corpus data alone, we do not really know whether the language user is "aware" that the verb-intensifier combination in question is also a realisation of a more abstract pattern or not. In order to empirically test this upward strengthening hypothesis, Hilpert (2015a: 140) suggests to "re-create actual historical developments, as observed in diachronic corpus data, with a model that represents a dynamically changing constructional network". One possibility is to apply a computational model that simulates events of language use and readily accommodates for highly complex networks containing multiple connections that are constantly changing, as described in van Trijp & Steels (2012).

As the present investigation mainly relies on frequency and semantic information that is directly observable in actual corpus data, we wish to remain careful with respect to cognitive implications of the diachronic shifts that were observed in the construction (see Ch6, §6.2.3.2 on the cognitive status of network representations and entrenchment). We mainly hope to have illustrated that, from a more linguistic point of view, multiple mechanisms – including conventionalisation and schematisation – have shaped and reshaped the taxonomic representation(s) of the constructional network of the intensifying fake reflexive resultative construction, which has resulted in a complex, dynamic constructional network.

5.5 Second interim conclusion

5.5.1 Expressivity, productivity and shifts in the constructional network

Rather than summarising the historical development of the intensifying fake reflexive resultative construction (which will be briefly recapitulated in Chapter 6, §6.1), this second interim conclusion will primarily highlight some of the specific changes and shifts that are interesting from a more theoretical point of view and will be of importance in Chapter 6, §6.2.

First of all, the specific semantic-pragmatic nature of the intensifying fake reflexive resultative construction, which was characterised in Chapter 2 as an "expressive" construction, may explain some of the changes observed in this chapter. The need for expressivity or "extravagance", which was said to be a driving force in the domain of linguistic intensification, may tie in with several of the (productivity) shifts the construction has undergone in the course of the last 200 years or so. At the maximum level of abstraction, the construction has undergone a general expansion with respect to both token frequency and verb and intensifier type frequencies. In particular, the increasing growth and diversification of the intensifier repertoire, as well as the linguistic creativity that is observed in present-day Dutch can plausibly be related to the expressive meaning component of the construction. If language users (in this case, journalists) strive to be noticed for their "linguistic prowess", this may explain why we often see new, creative intensifiers being introduced in the construction. That is, as several intensifiers increase in frequency and spread to an ever increasing pool of verb types, they may no longer be felt to be sufficiently strong or expressive in specific contexts and are prone to be replaced by a more extravagant alternative whenever the speaker feels the need to use a "special" means of intensification, consciously or unconsciously. Once they are introduced, some of these "new" intensifiers may also become increasingly frequent and extend their use to other verbs, which makes them vulnerable to losing their expressive force just like the intensifiers they had once replaced. This pragmatic wear-and-tear results in a constant cycle of innovation and renewal at the level of the [SUBJ V REFL INT] pattern (Stoffel 1901, Bolinger 1972, Partington 1993, Lorenz 2002, De Clerck & Colleman 2013, inter alia, see Ch2, §2.3). However, this process of renewal is not complete in the sense that existing intensifiers are simply replaced by new alternatives. The only clear example of an intensifier that was quite popular for several decades and that appears to have fallen out of use in Netherlandic Dutch (at least in SoNaR-NL) is *een aap* 'a monkey'. Other intensifiers have obviously decreased in frequency or retreated to particular collocations, but they have not disappeared entirely. The mere fact that the range of intensifiers has been consistently increasing over the past decades and that there are so many different intensifiers being used in the construction in present-day Dutch clearly

shows that multiple intensifiers can co-exist. As has already been shown by Ito & Tagliamonte (2003: 277) for English, this type of layering (if we borrow a term from grammaticalisation research) is typical for the use of intensifiers, as "old intensifiers do not fade way, they stick around for a very long time". If we recall Hoeksema's (2005, 2012) claim that "anything to do with degrees belongs to a part of the grammar where lexical parsimony is valued the least", the degree of variation attested in the intensifier slot should come as no surprise. In other words, expressivity appears to feed into creativity, which in turn may contribute to the productivity of the intensifier slot of the [SUBJ V REFL INT] pattern: the intensifier slot can be instantiated by an increasingly large and semantically diversified range of intensifiers.

Evidently, the intensifiers do not always co-exist in perfect harmony: the data contain some evidence of a certain competition at the level of individual intensifiers and the verb-intensifier combinations they enter into. For example, we found that the recently introduced intensifier *wezenloos* 'vacant/blank', which has only 66 total tokens, occurs with a wider range of verb types (23) and hapaxes (16) than a long-established, frequent intensifier like *dood* 'dead' (243 tokens, 17 types, 8 hapaxes). In fact, for several of the other highly frequent intensifiers as well, such as *kapot* 'broken', *rot* 'rotten' or *te pletter* 'to smithereens', the range of verbal collocates appears to have remained more or less constant in recent times, or even to have decreased. It appears that these intensifiers are primarily used with some of the more frequent verbs (e.g. *schrikken* 'to be startled', *lachen* 'to laugh', *zich ergeren* 'to be annoyed', etc.), but proportionally much less so with infrequent or hapax verbs, especially compared to other intensifiers like *suf* 'drowsy' or *wezenloos* 'vacant/blank'. This may suggest that they are developing into more neutral intensifiers that are primarily used in conventional contexts, but that are not necessarily the first choice when an intensifier is needed for a verb that does *not* enter into the fake reflexive resultative pattern all too often. The competition with other lexical items may have even caused formerly productive intensifiers to retreat to very specific collocations, as was attested for *wild* 'wild' and (perhaps) *blauw* 'blue'. In other cases, it appears that intensifiers which are introduced as part of specific collocations never really manage to extend their use beyond this collocation (e.g. *groen en geel* 'green and yellow', *een hoedje* 'a little hat', which are very much restricted to the combinations with *zich ergeren* 'to be annoyed' and *schrikken* 'to be startled', respectively). Nevertheless, the exact effects of this competition between intensifiers are somewhat unpredictable and it appears that not all intensifiers are equally "affected" (see also Ch6, §6.2.2). An interesting case in that regard is *suf* 'drowsy'. As it is one of the oldest intensifiers attested in our corpus (it was already attested in our corpus in the 1830s), we might expect it to be losing some ground to newer intensifiers in the more recent parts of the data set, much like was found to be the case for *dood* 'dead'. Quite to the contrary, however, *suf* has been steadily increasing its collocational range over the past 50 years and is currently by far the most flexible intensifier, at least in Netherlandic Dutch. Of course, given that *suf* 'drowsy' was limited

to a very specific semantic domain until the mid-20th Century, viz. the domain of mental activity verbs, its combinatorial flexibility is of a much more recent date than that of *dood* 'dead'. Unlike *dood* 'dead' or *kapot* 'broken', which are primarily used with frequent verbs, *suf* 'drowsy' appears to be the preferred or default intensifier for infrequent verb types, with a hapax count of no less than 35 in the 1990s.

In section 5.4, we investigated the implications of the observed expansion and conventionalisation on the hierarchic structure of the constructional network of the intensifying fake reflexive resultative construction. First of all, it is clear that the macro-construction [SUBJ V REFL INT] has increased its degree of schematicity, following Barðdal's (2008) proposal that the degree of schematicity is determined by the number of intermediate levels it overarches. Several verbs and intensifiers that used to be very infrequent and situated at the micro-construction level have increased their frequency and expanded their collocational range, giving rise to multiple intermediate subschemas at different levels of abstraction. However, at some point, the changes in the network may have no longer contributed to the schematicity of the most abstract node. Given the competition between intensifiers and the coming and going of creative hapax intensifiers, there is a lot of movement in the lower levels of the network, with specific constructs and micro-constructions emerging and falling out of use without necessarily affecting the overall network structure all that much. Although the schematic representation of the intensifying fake reflexive resultative construction at the most abstract level was definitely strengthened in the earlier periods, many of the recent reorganisations within the network may primarily have involved strengthening (or, occasionally, weakening) of low-level subschemas and micro-constructions (cf. §5.4.5 supra). We illustrated some of the most important reorganisations that have taken place in the network by providing a verb-centred and an intensifier-centred representation of the constructional network at four different intervals, viz. the 1890s, the 1940s, the 1970s and the 1990s. In the early days of the construction, each newly attested verb-intensifier combination enters the network at the bottom of the hierarchy, as a concrete construct. The language user may abstract away from the specific subjects and reflexive pronouns, giving rise to a micro-construction in which the SUBJ and REFL slots are left open but in which the verb and intensifier are lexically specified, viz. [SUBJ *specified verb* REFL *specified intensifier*]. As one of the two lexically specified items (or both) is/are extended to other items, it is possible that the language user recognises the combinatorial flexibility and forms a generalisation in which only the verb or the intensifier is specified. We saw that this already happened quite early on for the intensifier *dood* 'dead' and verbs like *lachen* 'to laugh' and *werken* 'to work'. If the items that occur in the open slot are found to form a coherent syntactic or semantic class, it is possible that an intermediate subschema is posited in which the open verb or intensifier slot is no longer lexically specified, but, importantly, is still semantically and/or syntactically specified. Only when types from multiple different classes have been attested in the slot can this restriction be lifted,

which gives rise to a higher-order subschema. A type increase within a certain subschema may indicate that an abstraction has been formed: once a subschema has been established, it may trigger the creation of new types. A clear example of this kind of semantic expansion was provided by *suf* 'drowsy', which was first confined to a set of mental activity verbs, before opening up its distribution to all kinds of different verb types. Of course, not all verbs and intensifiers actually give rise to a subschema: if their combinatorial flexibility stays extremely limited, they may be "stuck" on the micro-construction level. A case in point is *groen en geel* 'green and yellow', which has been around since the first half of the 20th Century but has virtually always exclusively combined with the verb *zich ergeren* 'to be annoyed' and is still restricted to that particular combination in present-day language. Language users may also come to perceive higher-order similarities between the different verb and intensifier types that are found in the construction. For instance, we mentioned that language users may recognise the semantic similarities between different intensifiers, e.g. the negatively connoted states or the diseases, and form productive semantic intensifier categories, such as [SUBJ V REFL $INT_{disease}$], at a higher level in the network. This is how the widening of semantic scope may come to influence the representation of the constructional network (cf. §5.1.1 and §5.1.2 supra). This continuous process of generalisation and abstraction has led to a very intricate network of subschemas and micro-constructions at different levels in the hierarchy. All the nodes are in one way or another licensed by a higher-order schema, but there is a lot of variation in the number of sublevels that one has to distinguish for specific verbs and intensifiers and some areas of the network may be more densely populated than others. Of course, specific nodes may also disappear from the network: if formerly productive verbs or intensifiers are losing ground and retreating to particular collocations, the overarching generalisation may lose strength and eventually disappear, as we saw for [SUBJ V REFL *wild*] and [SUBJ V REFL *blauw*]. In those cases, the other types were generally not semantically related to the more frequent item, which is unexpected if one were to consider them as analogical extensions but which makes sense if they are scattered relics of a formerly productive schema.

In sum, this chapter has provided a detailed picture of how the internal structure of the network has been reorganised. Importantly, it has shown that different mechanisms, such as shifts in productivity, analogical extensions and conventionalisation effects are at play at different hierarchic levels in the network. Only by looking at lower and intermediate levels of the network can we get a better understanding of the complex organisation of the constructional network at any given time.

5.5.2 Discussion: constructional variation and change in context

Having presented a detailed overview of the historical development and variation of the intensifying fake reflexive resultative construction, we briefly need to address to what extent the observed developments are potentially influenced by "environmental" or "contextual" changes. In a recent paper, Szmrecsanyi (2016: 154) argues that:

> fluctuating frequencies of grammatical variants are a function not only of changing grammars but are also conditioned by environmental changes in the textual habitat. So the crucial problem is that diachronically variable text frequencies often entangle environmental differences and grammatical changes.

In his case study on the genitive, he investigates the curious development of *s*-genitives in the Late Modern English period, which is often related to language-internal questions of (de)grammaticalisation of the genitive marker. He finds that the temporary frequency decrease of *s*-genitives around the turn of the 19th Century is concomitant with a more general decline of animate nouns, which are not coincidentally the preferred possessor category of *s*-genitives. This change in the distribution of animacy categories is related to a shift in the news genre, which has come to topicalise inanimate over animate entities. Singling out the animate category in the frequency development shows that the decrease in animate topics is indeed responsible for part of the frequency decrease of *s*-genitives. While the potential impact of environmental factors is widely recognised in variationist linguistics, Szmrecsanyi states that corpus linguists often heavily rely on frequencies without explicitly addressing the possible confounding contextual factors. It was mentioned in §5.1 that recent years have indeed seen an increasing interest in frequency changes as a worthy object of study in their own right, to the extent that Hilpert (2013: 16) puts frequency changes on a par with formal or semantic changes: "Constructional change selectively seizes a conventionalised form-meaning pair of a language, altering it in terms of its form, its function, any aspect of its frequency, its distribution in the linguistic community, or any combination of these". Of course, language cannot – and should not - be studied in a vacuum because language does not exist without actual language users interacting in a specific context (Curzan 2009: 1103). While it may often be difficult to disentangle actual language changes from other contextual changes, there are ways to at least partly overcome this problem. The solution offered by Szmrecsanyi (2016) is to complement the study of frequency changes with other aspects of use. While he finds that the decrease in animate possessors goes a long way in explaining the frequency dip in *s*-genitives in the 19th Century (cf. supra), it is less successful in predicting their revival in the 20th Century. In order to investigate whether this revival is truly a grammatical change, he uses mixed-effects binary logistic regression to measure the impact of language-internal factors and their interaction with the language-external

factor of time. It turns out that the effect of some language-internal conditioning factors does indeed change over time, which can be seen as evidence for grammar change.

Let us consider the implications of this discussion for the findings presented in this thesis. In Chapter 3, §3.1, we motivated why journalistic data do constitute a suitable genre for the type of investigation that is presented in this thesis, even if this might seem to go against intuitive impressions of a rather low degree of creativity in journalistic language. In addition to some important practical considerations that effectively ruled out other genres, there were also some other arguments in favour of journalistic data, the key point being that journalese, i.e. the language used in newspapers, is less "dry" or less formal than is sometimes assumed. While the trend towards a more informal and oral style in newspapers had already set in around the late 19th Century, the breakthrough of this "informalisation" or "conversationalisation" is generally situated around the 1960s. It was mentioned that this shift in style has provided journalists with more freedom to experiment or allowed them to be more creative with respect to their language use. In that regard, it is not implausible that the specific type of "expressive" or "intensifying" construction under investigation in this thesis was considered to be less acceptable or suitable in the more formal or "serious" newspaper style in the 19th Century. As newspapers gradually adopted a more informal type of language use, the genre may have become more accepting towards expressive constructions like the intensifying fake reflexive resultative construction (see, e.g., Mair 2006 on several examples of colloquialisation in English). Such a development may well be behind at least part of the frequency increase that is attested in the construction, although it cannot directly account for the peculiar rise-and-fall pattern that was attested in the frequency curves between the 1950s and 1980s. It might also have contributed to the increase and diversification of types, especially in the intensifier slot. It is not unlikely that several intensifiers were already used in the construction, but were deemed wholly inappropriate in journalistic texts in the older periods of the timespan under investigation here – think of the taboo disease intensifiers like *de pleuris* 'the pleurisy' or *de tering* 'the consumption' in particular. As the style of newspapers relaxed and became more informal, these intensifiers may gradually have become more acceptable in writing as well – which, note, would still constitute a constructional change, but one of pragmatic or stylistic expansion rather than actual emergence. It seems a bridge too far, however, to assume that all or most of the recently added intensifiers were already perfectly "acceptable" in the construction in the early 19th Century but only started showing up in our data set around the mid-20th Century because of the informalisation of the genre. If that really were the case, we would at least expect some trace of their use as intensifiers in a source like the *Woordenboek der Nederlandsche Taal*, which encompasses a variety of different (written) genres. In other words, even if the process of informalisation has increased the acceptability of some intensifiers in the newspaper genre, it does not provide an adequate

explanation for the introduction of all new intensifiers in the second half of the 20th Century.

The potential influence of contextual factors should also be taken into account when studying synchronic data. For example, Levshina et al. (2013) suggest that some of the observed tendencies in the variation between the causative verbs *doen* 'to do' and *laten* 'to let' in Belgian and Netherlandic Dutch "may be caused by country-specific peculiarities of the newspaper genres", but they do not further address what these peculiarities could be. To give an example from our data, we suggested that the higher frequency of the verb *rijden* 'to ride' in the Belgian variant of the construction could be partly explained by a (hypothesised) higher proportion of articles on cycling competitions in the Belgian newspapers. Still, given the large frequency discrepancy, it is quite unlikely that this is really the only explanatory factor.

It is important to note that the potential impact of genre changes mainly concerns changes in frequency (which may be related to changes in acceptability) and productivity of the verbs and intensifier slots at the maximum level of abstraction, i.e. the level of the schema [SUBJ V REFL INT]. One of the main arguments of this thesis is that it is often more informative to look at lower level of abstractions, focusing on the collocational behaviour of the individual elements that can be used in the open slots of the construction. On the basis of the covarying collexeme analyses and a multidimensional model for measuring productivity, we identified several different kinds of phenomena in the data that go beyond a widespread increase in frequency or variability. At the level of specific verbs and/or intensifiers, we found a lot of variation with respect to the moment at which they were introduced, the rate of their expansion and their pathways of change. For instance, we discussed a number of intensifiers that originally only occurred with verbs from a delineated semantic domain, a domain which was to a large extent determined by the lexical semantics of the intensifier. Over time, the semantic persistence effect weakened and the intensifiers were extended beyond this semantic domain. This specific kind of type expansion, which goes hand in hand with a relaxation of collocational constraints as a result of semantic bleaching, is not easily explained by the earlier-described changes in the genre.[73] Furthermore, if we look at specific verb-intensifier combinations in more detail, we find several shifts with respect to collocational preferences or conventional collocational patterns that present themselves as construction-specific, idiosyncratic conventionalisation effects. Although the history of the construction is broadly speaking characterised by expansion and increasing creativity, we also witness the emergence and obsolescence of several conventional patterns at different points in the timeline. It is

[73] This would, for example, imply that *suf* 'drowsy' was already used with this wide variety of verbs in the 19th Century as well, but that for some reason, it only occurred with the mental activity verbs in newspapers – which is rather unlikely.

possible that the established collocations find themselves in rivalry with more recently introduced verb-intensifier combinations. In the most extreme scenario, the elements of the collocations are ousted by the newer introductions to the extent that they are no longer (or barely) used in the construction at all. In other cases, the competition pushes some items which used to have a wider distribution back to extremely specific uses. Although the informalisation of newspapers may have contributed to the recent explosion of new intensifier types which lies at the basis of an increased competition, it offers no explanation for the sometimes peculiar interactions of these intensifiers with one another and with their verbal collocates.

To summarise, it would be a vast oversimplification to reduce all constructional changes to environmental factors. While we do not deny that changes in the genre do potentially play a role in the sense that they may have supported or even fuelled certain developments of the construction overall, they cannot explain the diversity of changes that were attested at the level of the specific verb or intensifier. We argue that it is the interaction between creativity or productivity and convention(alisation), which is determined by different kinds of frequency effects (like statistical preemption or low-level analogy), competition and (often construction-specific) semantic effects, which has determined the historical development and continues to govern the synchronic variation of the intensifying fake reflexive resultative construction. Of course, it would be interesting to compare the results presented in this thesis to diachronic data for other genres, but this would require the compilation of a stylistically differentiated diachronic corpus for Dutch (cf. Ch6, §6.3).

Chapter 6 Discussion and conclusion

6.1 Synchronic and diachronic variation in the intensifying fake reflexive resultative construction: a brief recapitulation

The investigation presented in this thesis has provided a detailed description of synchronic and diachronic variation in the intensifying fake reflexive resultative construction, both at the level of the abstract construction and at the level of specific verbs and intensifiers, as well as at intermediate levels. In this first section, we will summarise the main findings of Chapters 4 and 5, focusing primarily on the variation and change that was observed in the construction, without insisting on theoretical notions like productivity and constructional networks. The next section will elucidate how the results of this very specific case study can be interpreted within the larger theoretical framework of Diachronic Construction Grammar.

6.1.1 The intensifying fake reflexive resultative construction in present-day Dutch

In Chapter 4, we investigated the use of the intensifying fake reflexive resultative construction in present-day Netherlandic and Belgian Dutch. Overall, the construction is used more frequently and with a slightly greater variety of different verbs and intensifiers in the Netherlandic data, but it appears that the construction is characterised by a high degree of variability in both slots in present-day Dutch in general. While a large portion of the data set was accounted for by a relatively small number of highly frequent verbs and intensifiers in both national varieties – suggesting that the use of the construction is conventionalised to a considerable degree – the high proportion of hapax legomena also showed that the construction is productive and allows for a great deal of

creativity. The verb slot can be filled by virtually any verb that has some inherent aspect that can be intensified or boosted. The most frequently used verbs denote either a cognitive or emotional experience (e.g. *lachen* 'to laugh', *schrikken*, 'to be startled', *zich ergeren* 'to be annoyed', *zich vervelen* 'to be bored', *zich schamen* 'to be embarrassed'...) or a heavy physical activity (e.g. *werken* 'to work' or *lopen* 'to run'), suggesting that these are the verb classes that are especially prone to intensification by this specific construction. In addition, the construction is also compatible with communication verbs, consumption verbs, noise emission verbs and a wide range of other activity verbs. In fact, most of the infrequently attested verbs are not easily classified in any of the larger categories and are scattered across semantic space.

The intensifiers are also recruited from several semantic domains and belong to different syntactic types. Much like was found for the verbs, not all formal and semantic subcategories of intensifiers are equally well-represented in the data. The majority of the highly frequent intensifiers belong to the subcategory of adjectives originally denoting a negatively connoted state (e.g. *dood* 'dead', *kapot* 'broken', *rot* 'rotten', *suf* 'drowsy'), although there are also a couple of highly frequent prepositional intensifiers, most notably *te pletter* 'to smithereens' and *uit de naad* 'out of the seam' – which, like the highly frequent adjectives, also have a somewhat negative connotation. The most type frequent semantic subcategory is the category of the diseases, which, formally, is chiefly made up of NP intensifiers: most of these intensifiers are individually not very frequent, but the diversity of (real and fictitious) diseases is highly characteristic of this construction, e.g. *de pleuris* 'the pleurisy', *de tering* 'the consumption', *de klere* 'the cholera', *het schompes* 'fictitious disease', etc. A smaller category of NP+PP intensifiers involves an inalienable body-part or piece of clothing, e.g. *het vuur uit de sloffen* 'the fire out of the slippers', *de longen uit het lijf* 'the lungs out of the body'. It appears that many intensifiers frequently occurring in this construction have a more or less strong negative connotation in their original semantics, which may have paved the way for an intensifying meaning to arise. This is much less obviously the case for the category of the colour intensifiers (e.g. *blauw* 'blue', *groen en geel* 'green and yellow'), the use of which may have been inspired by existing symbolic associations between colours and certain emotions. Overall, there are only a couple of (both conventional and unconventional) intensifiers that cannot be categorised into one of the established semantic categories (e.g. *een hoedje* 'a little hat', *een slag in de rondte* 'a punch around', etc.). The differences between Netherlandic Dutch and Belgian Dutch are primarily situated at the level of the intensifier: it appears that speakers of Netherlandic and Belgian Dutch hold different preferences and may rely on different intensifiers in specific situations. Even though there is a substantial overlap of intensifiers occurring in both national varieties, some intensifiers are significantly more frequent in one of the two national varieties. In addition, there are a number of nationally exclusive intensifiers, i.e. intensifiers that are somewhat frequent in one variety (i.e. at least 10 occurrences) but are altogether absent in the other. The intensifiers typical for

Netherlandic Dutch are *een slag/slagen in de rondte* 'a punch/punches around', *wild* 'wild' and *ongans* 'unwell'; the exclusively Belgian Dutch intensifiers are *de ziel uit het lijf* 'the soul out of the body', *een aap* 'a monkey', *steendood* 'stone-dead', *een beroerte* 'a stroke', *zot* 'crazy' and *de naad uit het lijf* 'the seam out of the body'.

While we can consider the verb and the intensifier slots separately, a detailed study of the collocational patterns in the construction showed that there are important interactions between the two slots. Some verbs and intensifiers co-occur much more frequently than would be expected on the basis of their individual frequencies, and some even enter into near-exclusive associations with just one or two other items. Both national varieties share a number of such conventional collocations, e.g. *zich een hoedje schrikken* 'to startle oneself a little hat', *zich blauw betalen* 'to pay oneself blue', *zich suf piekeren* 'to worry oneself drowsy', *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers', *zich uit de naad werken* 'to work oneself out of the seam', etc. If we look at the collocational behaviour of the intensifiers that are highly typical for one of the two national varieties, we find that some of these are also part of conventional collocations, e.g. *zich groen en geel ergeren* 'to annoy oneself green and yellow', *zich wild schrikken* 'to startle oneself wild' and *zich wild ergeren* 'to annoy oneself wild' in Netherlandic Dutch or *zich een aap schrikken* 'to startle oneself a monkey' and *zich steendood vervelen* 'to bore oneself stone-dead' in Belgian Dutch. In those cases, we have argued that it may be more accurate to refer to different *idiomatic expressions* in Belgian and Netherlandic Dutch rather than to nationally exclusive or preferred intensifiers. On the whole, the similarities between Belgian and Netherlandic Dutch outweigh the differences. Nonetheless, the observed national preferences clearly indicate that speakers of Belgian versus Netherlandic Dutch have partly different sets of conventionalised uses of the construction. For example, the number one expression in Netherlandic Dutch, viz. *zich groen en geel ergeren* 'to annoy oneself green and yellow', may actually sound rather odd to a Belgian speaker. Conversely, any combination with the intensifier *de ziel uit het lijf* 'the soul out of the body', which has almost 100 occurrences in Belgian Dutch, may raise some eyebrows in a Netherlandic context.

In sum, the results suggest that the Dutch language user often does not randomly put together a verb and an intensifier and that the individual items display substantial differences with respect to their combinatorial flexibility and collocational preferences. The intensifying fake reflexive resultative construction in present-day Dutch displays an interesting combination of creativity and convention. In sections 6.2.2 and 6.2.3 we will further elucidate how these findings can deepen our understanding of the concept of productivity at different levels in the constructional network and the mechanisms which underlie this hierarchic network architecture.

6.1.2 Diachronic development of the intensifying fake reflexive resultative in Netherlandic Dutch (1830-1995)

Chapter 5 set out to trace the history of the construction since the early 19th Century. A closer look at the different frequency aspects of the construction has shown that the construction has gradually undergone a significant increase in overall frequency, as well as a considerable expansion of the observed types of both verbs and intensifiers. A crucial phase in the expansion of the construction seems to be situated around the 1930s, from which moment onwards the type and token frequency curves were found to increase more steeply. Although the curves run largely parallel, the increase in token frequency is much more exponential than the increase in (verb and intensifier) types, because the expansion of the construction is mainly carried by a number of highly frequent verbs and intensifiers. At the same time, the second half of the 20th Century displays some marked fluctuations, especially in the intensifier type frequency curve. Taking the rise-and-fall pattern under further scrutiny, we found that the 1950s and the 1970s were characterised by an influx of new hapax intensifiers, the large majority of which did not make it into the next decade. In fact, taking the hapaxes out of the equation gives a much more linear trend.

Zooming in on the developments of some individual verbs and intensifiers, we found that there appears to be a certain diachronic consistency in the kinds of verbs that are particularly suited for being used in the intensifying construction under investigation, whereas there has been much less continuity in the actual intensifiers used for boosting these verbal meanings over the past two centuries. While several of the verbs that are currently very frequent were already among the most prominent verbs in the late 19th Century as well, most of the frequently used intensifiers in present-day Dutch were only introduced in the second half of the 20th Century. The higher volatility of the intensifier slot may be related to the fact that the linguistic domain of intensification is characterised by constant innovation and renewal (see Ch5, §5.5.1 and §6.2.2 infra). If we widen our scope to include all infrequent verbs and intensifiers, both slots show clear signs of semantic expansion. The (still dominant) classes of experience verbs and physical activity verbs were already represented in the 19th Century, but we only start seeing some examples of consumption verbs, verbs of noise emission and communication verbs around the turn of the 20th Century. The second half of the 20th Century also sees all kinds of different activity verbs that do not easily fit into one of the larger semantic classes, joining the distribution. It is this large group of "other" or "general activity" verbs that is the main contributor to the type explosion in recent decades. With respect to the intensifiers, most of the items that were already present in the early and mid-19th Century belong to the class of negatively connoted states, the negative semantics of which have undoubtedly contributed to their being so readily used as intensifiers (cf. Chapter 2, §2.3). We also see a number of inalienable possession intensifiers showing up early in the

construction, but some of the other semantic classes that were established on the basis of present-day Dutch are not yet represented in the 19th Century. For instance, the frequent colour adjectives *blauw* 'blue' and *groen en geel* 'green and yellow' are introduced in the first half of the 20th Century, but only in the second half of the 20th Century do we start seeing analogical variations with other colour combinations. Most notably, the category of diseases, which is so prolific in present-day Dutch, is still largely underrepresented in the first half of the 20th Century. Aside from some general bodily ailments like *een bult* 'a hump', none of the (often informal) expressive terms for life-threatening diseases were attested yet.

In Chapter 4, we have seen that the individual verbs and intensifiers are involved in important interactions and that their distributional behaviour is to some extent determined by covariation or coselection. Given these collocational patterns, it is not always possible to disentangle the historical developments of individual verbs and intensifiers in the construction. The diachronic application of covarying collexeme analysis allows us to track the emergence (or obsolescence) of strong collocations in the intensifying fake reflexive resultative construction. There are only a couple of verb-intensifier combinations that have consistently featured in the top twenty of strongest collocations throughout the periods under investigation (i.e. *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' and *zich suf piekeren* 'to worry oneself drowsy'). The top collocations in the most recent decades contain several examples of verb-intensifier combinations that were only introduced in the mid-20th Century, but that have quickly become strongly conventionalised expressions. At the same time, there are several verbs and intensifiers that have expanded their collocational range over time. Importantly, as the verbs and intensifiers did not enter the construction at the same time, they differ with respect to the pace at which they have developed. The intensifiers *dood* 'dead' and *suf* 'drowsy', for example, have been attested since the early 19th Century but many of the other intensifiers that are quite productive in present-day Dutch (e.g. *rot* 'rotten' or *te pletter* 'to smithereens') are recent success stories which only entered the construction in the second half of the 20th Century. It appears that the increase in type frequency often goes hand in hand with a relaxation of collocational constraints over time, but there are also all-round intensifiers that already occurred with a semantically diverse set of verbs from the beginning, such as *dood* 'dead', *rot* 'rotten' and *een ongeluk* 'an accident': in increasing their type frequency, the different semantic classes that are already represented are further elaborated and become more dense, but there is no real semantic expansion. If an increase in type frequency is usually correlated with an increase in semantic range, by that same logic, a decrease in type frequency would be concomitant with a decrease in semantic scope, or a retreat to a delimited semantic domain. However, we found that the verbs and intensifiers that are losing ground in the construction are generally not really retreating to particular subregions of their former semantic range, but are rather shedding different individual senses that may be scattered

across semantic space. Moreover, there are cases of conventionalisation and fossilisation, in which former productive intensifiers are confined to specific collocations, that are not motivated by any obvious semantic reasons. The interplay between expansion and conventionalisation is one of the aspects that will be discussed in the next section, which takes a more theoretical approach and aims to illuminate how the detailed analysis of this specific construction can contribute to the developing field of Diachronic Construction Grammar.

6.2 Theoretical implications

The results presented in this thesis have shown that the intensifying fake reflexive resultative displays interesting variation both from a synchronic and a diachronic point of view. In present-day Dutch, the construction can be represented as an intricate constructional network (or an interaction between multiple network representations, cf. infra) with multiple intermediate levels of abstraction, displaying different degrees of productivity. Its expressive meaning component and the interaction between increasing expansion and conventionalisation make it an interesting object for studying recent constructional changes. Throughout the thesis, we have occasionally touched upon the broader significance of our observations in light of certain theoretical notions that were introduced in Chapter 2. In this section, we return to the theoretical framework that was set out at the beginning of this thesis in more detail, in order to see how our findings tie in with existing research within the domain of Diachronic Construction Grammar and how they can further add to our knowledge on the mechanisms or factors involved in constructional change and variation.

6.2.1 Constructional changes in the broader sense

Although it was not a central focus in the present investigation, we want to start by briefly addressing the distinction between constructionalisation and (other kinds of) constructional changes introduced by Traugott & Trousdale (2013), which continues to be one of the central issues in Diachronic Construction Grammar (see, e.g., Hilpert 2018 for recent critical discussion). In Chapters 2 and 3 (§2.2.2; §3.3.5), we suggested that the intensifying fake reflexive resultative construction is likely to have arisen out of the literal fake reflexive resultative construction, while at the same time providing a number of arguments in favour of viewing the intensifying construction as a construction in its own right, at least in present-day Dutch. For one, it was mentioned that the intensifying

construction clearly has different aspectual properties than the regular resultative construction, in that it encodes an activity and is found with time phrases expressing duration, which are incompatible with a resultative reading. In addition, there are numerous intensifiers in present-day Dutch that cannot be used as a resultative phrase and which, therefore, are highly unlikely to have been derived from a prior resultative use. At the same time, we admitted that it was not entirely clear at which point exactly the construction has entered the grammar as a new construction. Following Traugott & Trousdale's (2013: 22) observation that constructionalisation can be seen to have taken place "when constructs begin to be attested which could not have been fully sanctioned by pre-existing constructional types", we would have to assume that the intensifying fake reflexive resultative pattern had already become a construction in its own right before the 19th Century. The WNT and the Corpus Literair Nieuwnederlands (Geleyn 2016) contain several intensifying examples (e.g. *zich dood zweten*, *zich slap lachen, zich ziek lachen*...) from before 1800 for which an interpretation as examples of the literal fake reflexive resultative construction is highly unlikely. In light of these findings, it was explicitly stated that the focus of this investigation was not to track or document the constructionalisation of the intensifying fake reflexive resultative construction, but to study the different kinds of (constructional) changes it has undergone since the beginning of the 19th Century. However, even though there were already unambiguous intensifying uses of the pattern attested before the 19th Century, it is not entirely clear whether language users already recognised these specific instances as examples of a schematic [SUBJ V REFL INT] pattern. Given that the use of the "construction" was clearly still very limited at the time, it is possible that the already attested uses were viewed as a set of related constructional idioms, instead. Moreover, we noted that in addition to the examples with clear intensifying semantics, there were also some specific cases for which the literal-resultative semantics still lingered in the background. It is not implausible, then, that several of the early examples in our data set are examples of transitional stages (i.e. the "bridging contexts" in grammaticalisation, Heine 2002) and that the constructionalisation was not yet "complete".

Even though Traugott & Trousdale (2013) emphasise that (grammatical) constructionalisation is a gradual process, the fact that they posit a difference between pre- and post-constructionalisation constructional changes does suggest that there is some kind of discrete threshold (Hilpert 2018). While it may be insightful to distinguish between changes that are affecting just one aspect of a construction from changes that could lead to the emergence of a new construction, the distinction often appears to be difficult to make in practice. We could question whether we *always* need a strict dichotomy between constructional changes and constructionalisation in order to describe the diachronic development of a given pattern (also see Börjars et al. 2015). Our investigation has shown that it is possible to provide a detailed description of different kinds of changes affecting a construction (or pattern), without explicitly having to decide

on whether and when constructionalisation has taken place: the observed changes display an inherent gradualness and continuity that do not easily allow for a categorisation in terms of pre- or post-constructionalisation changes. In that regard, we have adopted Hilpert's (2013) broader definition of constructional change:

> Constructional change selectively seizes a conventionalized form-meaning pair of a language, altering it in terms of its form, its function, any aspect of its frequency, its distribution in the linguistic community, or any combination of these. (2013: 16)

In the remainder of this section, we will primarily focus on the two theoretical concepts that were at the centre of this thesis, viz. productivity and constructional networks, but we will also refer to the role of expressivity where relevant (see also Ch5, §5.5.1). First, it will be discussed in what way the present investigation has contributed to the study of the (diachronic) productivity of constructions. We will identify the factors that were found to play a role in determining productivity and reflect upon how the existing measures could be refined to provide a better empirical basis for studying quantitative and qualitative aspects of productivity. Related to the notion of productivity is the internal structure – and the reorganisations of that structure – of the constructional network. The second subsection will elucidate the mechanisms that come into play in (re)shaping the constructional network and discuss what we can gain from couching constructional variation and change in terms of a taxonomic Lexicality-Schematicity Hierarchy. We will finish by addressing the cognitive reality of such taxonomic representations, a topic that has recently become a subject of some debate in (Diachronic) Construction Grammar.

6.2.2 Constructional productivity in synchrony and diachrony

6.2.2.1 Studying productivity from a constructional perspective

In Chapter 2, it was observed that productivity as a linguistic phenomenon was first studied in the domain of morphology. In this morphological tradition, productivity was chiefly addressed from a synchronic perspective and was seen as a concept enabling comparison of the applicability of different word-formation rules (Aronoff 1976, 1983, Baayen 1989, 1990, 1992, 1993, 2009, Baayen & Lieber 1991, Plag et al. 1999, Anshen & Aronoff 1999). Studies on morphological productivity often focus on rivalling word-formation patterns, i.e. patterns that are essentially functionally equivalent, such as the deadjectival suffixes *–ity* and *–ness*. Evidently, there would not be much point in directly comparing the productivity of patterns that convey entirely different meanings or that belong to very different functional domains, as that difference in itself may influence the degree of productivity in important ways. In syntactic or constructional approaches to productivity, too, there have been several studies that rely on the notion of productivity

to compare different constructions that are in direct competition or operate within the same functional domain. Zeldes (2012) has illustrated how the empirical measures that were developed in the morphological tradition can also be used to compare the productivity of syntactic constructions (see Ch2, §2.1.2). Within the framework of Diachronic Construction Grammar, productivity also came into focus as an interesting phenomenon for diachronic research, in that certain changes within a construction are often concomitant with or indicative of changes in the productivity of that construction. In the final chapter of her monograph, Barðdal (2008) adds a diachronic dimension to her study on the productivity of case and argument structure constructions in Icelandic (e.g. Nom-Acc, Nom-Dat, Nom-Gen), by tracking how the productivity of those (sub)constructions has changed in Icelandic and several Germanic languages (see also Barðdal 2009, cf. Ch2, §2.1.3 for more details). In addition, there have been several studies which track the changes in productivity of very specific constructional patterns (e.g. Hilpert 2013 on the V-*ment* construction, Perek 2016a,b on the *hell*-construction and the *way*-construction). In this thesis we looked into productivity from a synchronic as well as a diachronic point of view. First, we compared the productivity of a construction and its multiple subschemas both within one national variety, viz. Netherlandic Dutch, and across the national varieties of Dutch. Second, we tracked the changes in productivity that this construction and its subschemas have undergone over the past two centuries in Netherlandic Dutch.

Our investigation is slightly different from the majority of existing work on productivity in two ways. First of all, most existing studies on productivity essentially deal with the productivity of one particular slot in the construction. In partly lexically specified patterns with only one open slot there is of course only one option, but there are several constructions with more than one open slot. In the case of argument structure constructions, the extensibility of the verb slot is often considered to provide a measure for the productivity of the entire construction, insofar as the construction is considered productive if it can attract novel verbs. As has already been pointed out by Zeldes (2012), however, depending on the research aim and the specific construction under investigation, other, i.e. non-verbal slots may be of equal or even greater interest. One possible way of dealing with multiple slots is to consider the productivity of one slot at a time in a kind of hierarchical fashion (Zeldes 2012: 125). This hierarchical or "nested" procedure follows the natural selectional process by which one slot is selected before another (e.g. the object is generally selected by the verb, rather than the other way around). While this is a valuable approach, the researcher still has to decide which slot he/she will give priority to, which may influence the results in important ways. We have illustrated that the two main open slots in the construction under investigation here (viz. the verb slot and the intensifier slot) were not always instantiated independently, but were instead often coselected (much like the comparative correlatives in Zeldes 2012). If we were to prioritise the verb slot over the intensifier slot, we risk ignoring the particular

nature of the construction and its historical development. Therefore, we have decided to evenly divide our attention over the two slots, evaluating the productivity of both slots at the same level of abstraction. Of course, there were some specific aspects for which one slot turned out to be much more informative than the other, but this is not something we could have known a priori, thus showing why a hierarchical approach may not always be the best option. It would be interesting to see whether, in existing studies on the productivity of other constructions, new insights could be gained from taking the perspective of another slot than the verb slot.

Related to the previous point is our decision to not focus exclusively or even primarily on productivity at the maximum level of schematicity. The results of the productivity analysis may not only be influenced by the open slot(s) that is/are selected, but also by the level of abstraction at which productivity is measured and compared. Within cognitive-functional linguistics, it is generally assumed that the productivity of a construction is determined by the most entrenched level (Clausner & Croft 1997). However, Barðdal (2008) contests this view because it would imply that the productivity of the construction is *confined* to that most entrenched level, which does not appear to be the case in her data. For the Nom-Dat construction, for example, the most type frequent (i.e. entrenched) subschema is the semantic caused-motion pattern, but the Nom-Dat construction can also be used outside of the semantic domain of caused-motion. Instead, she argues that it is not the most entrenched level which determines the degree of productivity but the construction's highest level of schematicity (2008: 45, 85). At the same time, she finds that even when a productive, higher-level schematic construction exists, a lot of specific verbs were assigned to the construction because of high similarity (or even synonymity) with a low-level verb-specific construction. Barðdal (2008) performed an experiment to find out which case and argument structure construction speakers of Icelandic would assign to nonce verbs. It appears that speakers assigned case and argument structure constructions *either* on the basis of a productive highly schematic construction *or* on the basis of a synonymous verb-specific construction. She even found evidence of one and the same speaker alternating between the two extension strategies within the same experiment (2008: 104). In a synchronic study on the generalisation strategies in a set of intensification patterns (cf. supra), Zeschel (2012) analyses the specific patterns from three different perspectives, viz. item-based generalisations, incipient productivity (i.e. low-level subschemas) and higher-order schemas, and also finds that it is crucial to take into account both item-specific information as well as more general semantic or other patterns in order to account for the productivity of these patterns. We will continue the discussion on productivity at different levels of abstraction in §6.2.2.1 below. Now that we have outlined some aspects and potential confounding factors that the researcher needs to take into account or be aware of when studying constructional productivity, we can turn to what our results have taught us on the factors that play a role in determining productivity and how these factors can be measured.

6.2.2.2 Measuring quantitative and qualtitative aspects of productivity

It appears that the productivity of a construction is driven by a number of factors that can be (directly or indirectly) observed or derived from corpus data. Generally speaking, the productivity of a construction is tied to the experience of language users with that very construction. That is, even though there is always some degree of unpredictability or arbitrariness involved (as we will show below), the extensibility of a pattern is to a large extent dependent upon the already experienced instances of that very pattern. While we cannot capture the precise experience of each individual language user, we have demonstrated that it is possible to get an abstract idea of the "general experience" of the linguistic community by combining two types of information, i.e. different types of frequencies and semantic aspects. These quantitative and qualitative aspects and the measures used to assess their impact on productivity will be discussed in the following paragraphs.

(a) Quantitative constructional productivity

Within the domain of quantitative productivity, three types of frequency information play a role in determining the productivity of a construction, viz. type frequency, token frequency and hapax count. We will discuss how these frequencies relate to productivity and how their contribution can be measured in actual (synchronic and diachronic) corpus data. In the traditional view on productivity, a high type frequency is often directly related to a high degree of productivity (Bybee 1985, 1995, Goldberg 1995, Bybee & Thompson 1997, Clausner & Croft 1997, also see one of the assumptions of the productivity model in Barðdal 2008, and one of the factors in Suttle & Goldberg 2011). It makes sense that the fact that a specific (sub)schema is already used with a high number of types may lead the language user to assume that it can also be extended to more types. Our data sets show that certain specific subschemas have become increasingly type frequent over the past two centuries. In those cases, the increasing type frequency may indicate that speakers have grown more and more confident of the extensibility of that subschema. Baayen (Baayen & Lieber 1991, Baayen 1992, 1993), however, argues that a high type frequency can only tell us that the construction must have been productive at some point – i.e. it gives an indication of the "realised productivity" or "past achievement" of a construction (cf. Ch2, §2.1.2 and Ch4, §4.3) –, but that it does not necessarily entail that the construction can still be used to coin new instances (see also Bauer 2001: 48, who similarly argues that "type frequency is the result of past productivity rather than an indication of present productivity"). If a specific subschema has increased its type frequency compared to previous periods, it has increased its *extent of use*, but the high(er) type frequency is not enough to warrant its extensibility to new types. Indeed, there were a number of subschemas in our data which, after a sudden type explosion, have afterwards remained at virtually the same level of type frequency in the following

decades. That is, while subschemas may currently have a relatively high type frequency, this only shows that the schemas in question were once able to attract new types, even if they no longer do so at present. The present investigation has thus confirmed the general importance of type frequency, but it has also shown that it is hard to pin down the exact implications of type frequency for productivity or the exact aspect of productivity it measures.

As an alternative measure for gauging the *extensibility* of a construction, Baayen has put forward the hapax-token ratio, i.e. the potential productivity measure $\mathscr{P}$. This measure is based on a combination of (i) the assumption that a high number of hapax legomena contribute to productivity and (ii) the traditional view that highly frequent tokens detract from productivity (Bybee 1985, 1995, Bybee & Thompson 1997, Clausner & Croft 1997). In other words, a pattern is considered extensible if it is not primarily used in highly frequent combinations but can appear with many different infrequent types. It is not clear, however, whether the proposed hapax-token ratio is the best measure to include the respective roles of hapax count and token frequency. One possible issue is that the ratio does not provide any insights into the way in which the tokens are distributed across the different types. This is problematic because a limited number of highly frequent tokens may artificially deflate the ratio and obscure the role of the hapaxes. In addition, there are some practical considerations to take into account. It was repeatedly stated in Chapters 4 and 5 (§4.3.1.1 and §5.3.1) that the potential productivity measure is highly sensitive to sample size, and that it is therefore far from ideal to use the measure for categories which display very different token frequencies. The variable-corpus approach provides a way of dealing with this, but for certain infrequent constructions, the largest shared sample size may be too low to yield any interpretable results. As an alternative, we suggested the hapax-type ratio, which measures the proportion of hapaxes to the total number of types rather than to the total number of tokens. This ratio is less sensitive to the influence of high-frequency tokens, but it has drawbacks of its own: it is not particularly useful for extremely low type frequency constructions, the extensibility of which may be overestimated by the ratio (e.g. if 2 out of 3 types are hapaxes, we get a rather high ratio of 0.67). Given that the studies on morphological productivity primarily take a synchronic perspective, it is also unclear how these ratios can be applied to and interpreted in diachronic data. First of all, the diachronic application of the hapax-token ratio suffers from the same practical issues as its synchronic application: in order to apply the measure of potential productivity to diachronic data, the samples for the sequential periods (ideally) are to be kept constant in size. Second, we can cast some doubt on the extent to which the potential productivity truly measures the extensibility of a construction in the long term. The original definition of the potential productivity measure is that it assesses the likelihood of retrieving a previously unattested type as the sample size is increased. We could theoretically extend this definition to linear time, in that it also assesses the likelihood of being extended to a

previously unattested type in the (immediate/near) future. Of course, if the $\mathscr{P}$-value is gradually increasing over an extended period of time, this can be interpreted as an increasing degree of productivity and, conversely, a steadily decreasing $\mathscr{P}$-score could be said to indicate that the construction is on its way to becoming unproductive (an example of the latter is discussed in Schneider-Wiejowski 2009, who find that in Swiss German the suffix *-sal* has gradually decreased its potential productivity and eventually became unproductive). However, the ratio in itself does not really have any "predictive power", in the sense that we cannot foresee whether (and how many) new types will be added in the future based on the $\mathscr{P}$-score in a given period – especially when dealing with periods that span an entire decade or even multiple decades. The comparison of potential productivity scores in Chapter 5 showed that a high $\mathscr{P}$-score is not necessarily followed by a type or hapax increase, even though a high ratio should indicate a high degree of extensibility. What is more, we have even found sudden decreases in type and hapax frequencies, despite a high $\mathscr{P}$-score in the previous period.

Some of the objections lodged against both ratios may be partially met by including a concrete measure that captures the type-token distribution of the construction, as a highly skewed, Zipfian distribution has been said to promote the extensibility of a construction (Zeschel 2012, Zeldes 2012, Gries 2012). We explored to what extent the relative entropy measure could serve this purpose, but after applying the measure to our synchronic data, we had to conclude that its results were not unequivocal and we no longer used it in the remainder of the thesis. In a way, the idea that a Zipfian distribution contributes to productivity is also based on the assumption that productive constructions are characterised by a high number of infrequent tokens (long right tail). This particular shape of distribution also suggests that a couple of highly frequent tokens are not necessarily detrimental to the productivity of a construction. Indeed, our data contained several subschemas that were characterised by both a number of highly frequent combinations and a considerable set of infrequent combinations, which prompts us to reconsider the role of high token frequency in productivity.

It appears that the effect of token frequency in productivity is somewhat ambiguous and not always easy to predict. On the one hand, the existence of a highly token frequent specific verb-intensifier combination may prevent the language user from extending one or both of the individual elements to other items beyond that collocation. This is the traditional view on the role of token frequency in productivity: highly token frequent instances of a construction are hypothesised to be accessed and analysed as a whole, rather than as on-the-fly productive instantiations of a productive pattern. The process by which the existence of a conventional collocation may block other possible combinations with those individual elements is captured by the notion of statistical preemption (Boyd & Goldberg 2011, Perek & Goldberg 2017). Our data contained several conventional collocations which consisted of at least one element that was not (or only very sporadically) used outside of that collocation, even though other potential

combinations are not strictly impossible or ungrammatical per se. On the other hand, a highly frequent combination may come to serve as a model for further extensions which in turn may give rise to a partially productive subschema. Barðdal (2008) gives several examples of case and argument structure constructions being assigned to new verbs on the basis of a high degree of similarity to just one existing verb, as is predicted by the bottom extreme end of the productivity cline (cf. infra). Token frequency is a relevant factor for the speaker's choice of model items in lower-level extensions, in that higher frequency items are more likely to be used as models than lower frequency items (see also De Smet 2016 on how token frequent, conventional expressions may have a better chance at triggering innovations than infrequent expressions). In addition, language users may replace one element of a highly frequent collocation or "fixed expression" with another element that is highly similar in form or in meaning. McGlone et al. (1994: 169-170) find that it is possible to substitute words in idioms and create new idioms, the meaning of which is based on the meaning of the original idiom and the meaning of the substitute words – a phenomenon for which they use the term "semantic productivity". Indeed, Zeschel (2012: 25) also noted that some so-called fixed expressions are "pre-assembled holistic units that are not assembled from scratch [but] not necessarily frozen in the sense of 'not tolerating any lexical subtitution' whatsoever". If such analogical variations or extensions are repeated multiple times, they may over time give rise to a more abstract schema that generalises over the specific lexical element. Once this abstract schema has been established, even more types can join the range of potential slot fillers, as a result of which the schema may become even more productive. Our diachronic analysis also provided several examples of present-day (partially) productive subschemas that have originally arisen out of one fixed, conventional collocation. Often, the collocation is still frequently used in present-day Dutch, next to a number of infrequent combinations. Indeed, this is the typical Zipfian distribution that was said to be characteristic for productive constructions. Based on these findings, it is not easy to pinpoint what the exact effect of token frequency will be: a highly conventional collocation may in fact have a preemptive effect, but it may just as well come to serve as a model for analogical variations and extensions. This also makes it difficult to capture the possible impact of token frequency with just one productivity measure. While the potential productivity measure is based on the idea that token frequency has a negative impact on token frequency, the positive impact of token frequency on productivity is better accounted for in the constructional model by Barðdal (2008), which will be discussed below.

(b) Qualitative constructional productivity

In addition to the quantitative aspects, we have also shown that the productivity of a construction is determined by its domain of application, which, in turn, is generally determined by its semantics. Suttle & Goldberg (2011) hypothesise that a construction

with a high degree of variability among its types is considered to be generally applicable by speakers and therefore likely to be extended to new types. At the same time, they suggest that a high variability may have a "dampening" effect, if most of the already attested types are very dissimilar from a specific target coinage. That is, in addition to the variability of the construction, the similarity between the already attested types and the new coinage is also of importance. A new type will be judged as more acceptable if it is highly similar to (one of) the types that are already used in the construction (see also Zeldes 2012: 185, who argues that new coinages of a schema reflect the semantic distribution of the already established types). In their experiment, they found important interactions between variability and type frequency, as well as between variability and similarity. The variability among the types was especially found to be an important factor when the type frequency was high and the similarity between the new coinage and the already attested instances was moderate (see Ch2, 2.1.2.2 for more details). Suttle & Goldberg (2011) calculated variability and similarity on the basis of Latent Semantic Analysis and Levin's verb classes, but as our corpus did not meet all the requirements for applying distributional semantic methods, we did not use these measures (see Ch4, §4.4 for discussion). In order to address the role of semantics in productivity, we have primarily relied on the constructional model of productivity suggested by Barðdal (2008). Within this model, the productivity of a construction is a function of its type frequency, semantic coherence, and the inverse correlation between the two. The inverse correlation is captured by a cline, where the most productive constructions are situated at the upper left corner (i.e. those constructions with high type frequency and little coherence between their types). More to the bottom right of the cline are the constructions that have a lower type frequency but the types of which display a high degree of semantic coherence. In the latter case, the construction is somewhat productive within a delineated semantic domain. If the lower type frequency is not supported by semantic coherence, the construction is situated in the left area underneath the cline and is expected to barely show any signs of productivity. At the most extreme bottom end of the productivity cline, we may expect to find extensions on the basis of similarity with one highly frequent type (i.e. analogy). Crucially, there is no qualitative or ontological difference between productivity and analogy, there are only different *degrees* of productivity. The inverse correlation also makes certain predictions about the diachronic productivity development of (rivalling) constructions that deserve some closer inspection. High type frequency constructions are expected to gain even more types over time, simultaneously widening their semantic scope even further. In other words, an upward shift on the cline means that an increase in type frequency is concomitant with an increase in semantic variability or a decrease in semantic coherence (e.g. through the relaxation of semantic constraints). This often happens at the cost of a rivalling lower type frequency construction, which is expected to fall into disuse (i.e. drop off the productivity cline). If the lower type frequency construction is reinforced by a high token

frequency, however, it may be preserved in a number of highly frequent lexically-filled instances. The extensibility of a low type frequency construction may be safeguarded if the decrease in type frequency is met with an increase in semantic coherence, in which case the construction shifts downwards to a lower point on the cline rather than falling off of it entirely. This was found to be the case for the Dative subject construction in Icelandic (Barðdal 2008, 2011). While the verb types were almost equally distributed across two semantic verb classes, viz. the experience-based predicates and the happenstance predicates in Old Icelandic, the Modern Icelandic Dative subject construction is primarily centred on the experience-based predicates. In other words, by dropping a number of verb types, the Dative subject construction has increased its semantic coherence and is still being extended to new verbs from that particular semantic domain. A rather similar development is found in the English ditransitive construction. While the ditransitive construction has decreased its semantic scope, it retains some degree of productivity because, in shedding some of the more marginal subsenses, its types have become more semantically coherent (Colleman & De Clerck 2011).

At first blush, the data presented in our investigation largely corroborate the hypotheses that are at the basis of the constructional theory of productivity as described above. There were a number of subschemas that could rather straightforwardly be positioned somewhere on the cline on the basis of their type frequency and semantic coherence. Our data contained examples of highly type frequent, semantically all-round subpatterns that could be situated at the top end of the cline, as well as highly token frequent instances that have served or could potentially serve as models for analogy at the bottom end of the cline. A number of partially productive subschemas with moderate type frequency and varying degrees of semantic coherence were positioned at different intervals in between the two extremes. Based on these findings, we concur that there is no discrete boundary between analogy and productivity, and, accordingly, that there is no qualitative difference between low-level or high-level extensions (cf. §6.2.3.1 infra). However, we ran into some trouble when trying to determine the relative positions of a number of specific subschemas on the cline. For example, some semantically coherent schemas turned out to be more type frequent than subschemas that were not subject to any obvious semantic constraints. Our data contains several local islands of productivity that showed very little internal coherence, thus casting some doubt upon the claim that "whether or not a low type frequency construction is productive depends *entirely* on the semantic coherence found between the types occurring in the relevant construction" (Barðdal 2008: 167, emphasis added). Even though Barðdal's constructional model of productivity has yielded good results for some argument structure constructions, these findings suggest that the strict linearity of the cline that embodies the inverse correlation may have to be somewhat relaxed. Of course, if we truly want to compare the varying degrees of productivity of different (sub)constructions, we would have to find a way to

determine their exact (X,Y) positions in the productivity plane (see Ch4, §4.3.1 and §6.3 on how to operationalise the model in a more systematic fashion).

The diachronic data presented in Chapter 5 contained (subtle) indications of competition between existing subschemas, suggesting that an exchange of types may have taken (and may still be taking) place between some of the lower-level subschemas. These subschemas are situated at the same level of abstraction and are thus in a kind of paradigmatic relationship (cf. 6.2.3.1 infra), which makes them highly sensitive to competition. Given the expressive meaning component of the construction and the notorious pragmatic wear-and-tear within the domain of intensification, the competition between the rivalling subschemas may be even more fierce than for other, more neutral constructions (Stoffel 1901, Robertson & Cassidy 1954, Bolinger 1972, Lorenz 2002, Norde et al. 2014, inter alia; cf. Ch2, §2.3). As predicted by the model, some highly productive constructions have drastically expanded their collocational range, whereas some erstwhile (mildly) productive subschemas have retreated – or are in the process of retreating – to lexically-specific collocations and others have fallen out of use entirely. Perhaps also related to this form of competition are the observed shifts in collocational preferences of specific subschemas, which could be indicative of collocational specialisation or differentiation. Some of the findings presented in Chapters 4 and 5 seem to suggest, however, that the success of one subschema does not *necessarily* happen at the expense of another subschema. There are several productive subschemas that occupy the same semantic domain and have shared a lot of specific types for an extended period of time, without any clear signs of one subschema "pulling away" types from the other. Although it has been assumed that languages tend to avoid synonymous constructions or that synonymy should be "eliminated" (Barðdal 2008: 145, 167), it may in fact be advantageous to have several forms for the same meaning (Traugott 2008b). If functional overlap is something that needs to be "solved", it is not clear why it emerges in the first place, or why the overlap is often sustained over an extended period of time. Indeed, the fact that new forms are constantly being introduced into the construction (cf. supra), thus *adding* to the already existing overlap and competition, does not mesh well with the idea that functional overlap should generally be avoided. De Smet et al. (2018) have therefore questioned the idea that the existence of functionally overlapping forms should only lead to two possible scenarios, viz. substitution (one form replaces the other) or differentiation (the two forms continue to co-exist in functional niches). They suggest that an alternative possibility is for items displaying a certain degree of functional overlap to become even more similar under specific circumstances, as a natural result of analogy. This kind of attraction may even facilitate parallel developments of functionally equivalent items. This is exemplified by some degree modifiers in Spanish, as discussed by Aaron (2016): once the first form *altamente* 'highly' had gradually developed into a degree modifier, the forms *extraordinariamente* 'extraordinarily' and *extremadamente* 'extremely' were able to adopt a degree function quite abruptly. For English, Margerie

(2011) discusses how the resultative phrase *to death* developed into a degree modifier over the course of several centuries (and is still primarily used as a resultative phrase in present-day English). As soon as the degree modifier constructions [NP_1 V NP_2 *to death*]/[NP BE ADJ *to death*] were established, other formally similar phrases like [NP_1 V NP_2 *rigid/stiff/silly*] received a degree meaning via analogy, without having to go through a long series of reanalyses (Margerie 2013). A very similar development is found in the intensifying fake reflexive resultative construction under investigation here. It was mentioned in Chapter 3 that the intensifying construction has originally developed from the literal fake reflexive resultative construction. Although we do not have data for the very first intensifying examples in our corpus, the examples in the WNT suggest that *dood* 'dead' may well have been the first resultative phrase to have developed an intensifying function. It is quite plausible that the other negatively connoted adjectives (*suf* 'drowsy', *ziek* 'sick', *kapot* 'broken', etc.) which came to function as intensifiers, did not have to go through the same development as *dood* 'dead' – even though they were also used as resultative phrases –, but were able to enter the construction more directly. This is definitely the case for the many intensifiers that were introduced in the construction without ever having a prior resultative use at all. Of course, it is also possible that certain similar forms continue to co-exist without undergoing any substantial changes at all. The tug-of-war between substitution and stability and attraction and differentiation makes it hard to predict what the "outcome" will be if the functional overlap is considered in isolation. Moreover, it was mentioned that the principle of economy may not be particularly important in specific semantic domains, such as intensification (Hoeksema 2005, 2012). If we take together all our findings, it appears that the expressive meaning component of the construction under investigation here does not only give rise to competition and renewal, but also to lexical profusion and layering within its specific domain of application (see also Ch5, §5.5.1). Our results are quite similar to those of D'Arcy (2015), whose study of a number of degree modifiers in English showed that the history of these intensifiers was not only characterised by "waves of recycling and renewal" (2015: 484), but also by longer periods of stability or "stasis". Ito & Tagliamonte (2003: 277), too, speak of both constant change or innovation and "extraordinary continuity" in the domain of English intensifiers, as they find that "old intensifiers do not fade away; they stick around for a very long time". While we observed a lot of diachronic fluctuation in the intensifier slot, it appears that, overall, the top five to ten of most frequently used intensifiers has stayed remarkably stable between the 1980s and present-day Dutch (and some of these intensifiers, such as *dood* 'dead' and *suf* 'drowsy' have been used in the construction since at least the early 19th Century). The renewal and innovation primarily seems to be taking place among the more infrequent intensifiers. Recently, some of the well-established intensifiers do appear to be losing some ground (cf. the subtle indications of competition discussed above), but their position among the most prominent intensifiers is safeguarded for now.

In general, there are several other qualitative factors that may impact the extensibility of low-level patterns in unexpected ways, sometimes giving rise to specific instances of a construction that are not easy to make sense of within the suggested multidimensional model of productivity. We saw that the immediate textual context may have some influence on the specific intensifier or verb-intensifier combinations that are selected. This could indicate that the productivity of a (sub)construction may be sensitive to (unconscious) priming effects, but the language user can also consciously opt for a specific combination in order to establish a partial parallelism with an expression that precedes or follows in the same sentence.[74] In other cases, a specific item may be selected or even newly formed to provide a better fit with the context or to create a play on words with other elements in the clause. In addition, language users may deliberately push the limits of existing conventions, such as conventional collocations or semantic constraints, in order to create a special rhetorical effect or to draw attention from the hearer/reader. These motivations are even more pertinent in the current investigation, as we are dealing with an expressive construction that is inherently prone to creativity (cf. Ch5, §5.5.1). Crucially, while the frequency aspects and semantic factors that were discussed earlier generally apply to the productivity of a construction at the level of the "linguistic community", the effects discussed here are often very context-specific or speaker-dependent and therefore hard to account for in a systematic fashion. For us, as linguists, it is not always easy to fathom what motivates an individual language user in producing a specific (especially non-conventional) instance of a construction (see Barðdal 2008: 93, 105 on how individual speakers may opt for other extension strategies and assign different case and argument structure constructions to the same verb). Given their idiosyncratic nature, the impact of such factors is relatively limited and they are unlikely to influence or skew the global productivity of a (sub)schema in any fundamental ways. Still, it is not entirely impossible that some of these low-level, idiosyncratic motivations could be picked up by other speakers and over time give rise to new productive schemas. Zeldes (2012: 228) formulates this as follows: "It is conceivable that [...] there is a certain 'critical mass' of formations (which speakers assume to be common ground with their interlocutors) beyond which creative extensions suddenly get off the ground in different semantic directions more freely".

To conclude, the fact that both quantitative and qualitative factors appear to have an important influence on the productivity of a construction clearly calls for a multidimensional model when measuring the synchronic or diachronic productivity of

[74] We saw, for example, that a language user may opt to use the same intensifier for all verbs in the same clause, even if some of these verbs are not conventionally used with that intensifier, e.g. *zich blauw werken* 'to work oneself blue' was used because *zich blauw betalen* 'to pay oneself blue' followed later in the same sentence and *zich uit de naad zingen* 'to sing oneself out of the seam' was coordinated with *zich uit de naad swingen* 'to swing oneself out of the seam'.

any construction. In this investigation, we have combined the frequency-based measures that were developed by Baayen and colleagues and the constructional productivity model by Barðdal (2008). We hope to have shown that this multidimensional model of productivity already goes a long way towards measuring and comparing the different aspects of productivity of schemas at different levels in the hierarchy. At the same time, there is still some uncertainty with respect to the exact interpretation of some of the measures and some areas of the model need further operationalisation.

6.2.3 A dynamic constructional network

From a usage-based point of view, the representation of the constructional network (at a given point in time) is built up from the bottom, starting with constructs or micro-constructions at the lowest level and further abstracting upwards if the data support the existence of more abstract (sub)schemas. This descriptive-linguistic usage-based approach to network representation reflects the historical development of the construction itself. Many schematic constructions have arisen out of specific, fully specified instances in actual language use, through repeated use and variation. In addition, it has been argued that the construct, i.e. the lowest level in the network, is the locus of innovation (Traugott & Trousdale 2013: 39). We elucidate how the taxonomic structure of the network allows us to interpret and clarify certain aspects of constructional variation and change. The first paragraph below reflects on how the Lexicality-Schematicity Hierarchy can account for the often complex nature (and development) of certain constructions by positing schemas at different levels of granularity or abstraction in the constructional hierarchy – which display different degrees of productivity and are subject to different kinds of (semantic) constraints. As several of the linguistic phenomena that will be discussed tie in with issues of the cognitive representation of constructions, we will address some of the cognitive implications of the network representation in the second part of this paragraph.

6.2.3.1 The network as a taxonomic hierarchy from schematicity to specificity

(a) A mix of productivity and idiosyncrasy

In this thesis, it was argued that the constructional network takes the shape of a hierarchy from lexically specified instances at the bottom to increasingly more abstract schemas at the top, a structure which has come to be known as the Lexicality-Schematicity Hierarchy (Croft 2003, Barðdal et al. 2011, Barðdal & Gildea 2015, Colleman 2015). Although it could be objected that this introduces some amount of redundancy in the description, this investigation has shown that the Lexicality-Schematicity Hierarchy offers a way of dealing with the combination of high-level productivity, intermediate levels with varying

degrees of productivity, and low-level idiosyncrasies that was displayed by the intensifying fake reflexive resultative in present-day Dutch. Of course, this situation is not unique to this construction, as several constructions can be said to subsume low-level "idioms" as well as subconstructions at various levels of abstraction that are subject to different types of constraints that do not operate on the construction as a whole. Within the framework of construction grammar, several studies have relied on some kind of specificity-schematicity hierarchy to account for the synchronic variation within specific argument structure constructions (Croft 2003, Barðdal 2006, 2008, Iwata 2008, Barðdal et al. 2011, Perek 2015, inter alia; see Ch2, §2.1.3 for more details on some of these case studies). In traditional rule-based or formal frameworks, this kind of synchronic variation in the range of application of a construction is sometimes considered to be "random", because it is not easily accounted for by a general syntactic rule. That is, in formal frameworks, any instance that is not fully explained by more general rules or that shows decidedly idiosyncratic properties is not strictly speaking part of the grammar. There are a number of studies which explicitly show that a network approach is preferable to a standard generative, rule-based system for any kind of pattern that shows a mixture of productivity and idiosyncrasy. A first example from morphology comes from Booij (2010a). It is often the case that a specific affix contributes different meanings depending on the type of base to which it is attached. In formal frameworks, one would have to posit multiple different word-formation rules for each of these senses. The Lexicality-Schematicity Hierarchy makes it possible to posit a general constructional schema specifying only the affix at the highest level of schematicity, while capturing the different possible contributions of the affix in (semantically or syntactically) specified subschemas and idiosyncratic formations at lower levels in the hierarchy (cf. the "hierarchical lexicon" in Booij 2010a: 77-80). Another example is presented by Jackendoff (2008), who finds that the Noun-Preposition-Noun pattern (e.g. *day by day, time after time*, etc.) covers several productive subpatterns with the prepositions *by*, *for*, *after* and *(up)on*, a more semantically constrained pattern with the preposition *to* and a collection of idioms like *tongue in cheek* or *hand over fist*. Jackendoff (2008, 2013) aims to account for the co-existence of those different kinds of subpatterns without having to relegate the semiproductive patterns or idioms to a domain outside of the grammar. He comes to the conclusion that "the distribution of NPN vividly illustrates the *continuity* between idiosyncrasy, semiproductivity, and full productivity argued for by construction grammar and related approaches" (Jackendoff 2008: 27, emphasis added). An interesting notion in this context is "pockets of productivity", which refers to the fact that there may be productive low-level patterns within a possibly less productive more general construction. Cappelle (2014) argues that the Body-Part-*Off* construction (as illustrated in the examples *He worked his butt off* or *She sang her heart out*) is a pocket of productivity within the caused-motion construction (which in itself is assumed *not* to be very

productive).[75] The BPOC itself consists of several subpatterns displaying varying degrees of productivity and a number of highly frequent conventional collocations (see Ch2, §2.2.1 for some examples). A more in-depth look prompts Cappelle to suggest that the BPOC "may best be analysed in terms of high-frequency learned instances and some creative extensions of these" (2014: 252). Interestingly, he proposes that the same analysis applies to the Dutch intensifying double-object cases, as in, e.g., *Hij schrikt zich een hoedje* 'he startles himself a little hat' or *Ze lacht zich een breuk* 'she laughs herself a fracture'. However, within this investigation, the intensifying double-object cases (i.e. the instances with NP intensifiers) are considered a subtype within the much broader network of the intensifying fake reflexive resultative construction. Much like the BPOC, the Dutch construction displays clear signs of conventionalised combinations and partial productivity at intermediate levels, but, as will be further discussed below, there are also indications that the pattern is productive at the maximum level of schematicity. Within the constructional network structure that is proposed here, all instances of a pattern – regardless of their level of granularity or their individual frequency – are included in the network. It is not necessary, then, to provide a separate explanation for the so-called conventional collocations (or "fixed expressions" or "idioms") or to store them separately in the lexicon. Even if those conventional collocations do have some special status within the constructional network (cf. §6.2.3.2 on the cognitive implications of this claim), the abstract schema is still immanent in their use and they are taxonomically motivated by higher-order subschemas, from which they inherit certain formal or semantic properties.

From a diachronic point of view, as well, there may be different kinds of changes happening within the same construction that upon first impression seem hard to reconcile. The overall frequency increase and the collocational expansion to new verbs that was observed at the maximum level of schematicity, rather straightforwardly qualify as diffusional changes (De Smet 2013) and also perfectly fit in with the post-constructionalisation constructional changes as discussed by Traugott & Trousdale (2013) (though cf. §6.2.1 supra on why we prefer not to use that term). At the lower levels of the network, there are several developments that blur the boundaries between

[75] In fact, Cappelle (2014) here refers to Kay (2013), according to whom the caused-motion construction is not a proper "construction" at all. Kay reserves the term "construction" only for fully general, productive patterns, the distribution of which is not subject to any constraints. Patterns like the caused-motion construction, the application of which is constrained in important ways, are referred to as "patterns of coining". Given that novel instances of these patterns of coining can be created with relative ease, it may seem as if they are productive, but Kay argues that the new coinages are created through processes of analogy and should not be interpreted as instances of a genuinely productive construction. However, we have already repeatedly stated that we do not assume a qualitative difference between extensions on the basis of low-level analogy or high-level productivity (following Barðdal 2008). Moreover, the fact that there are certain restrictions on the applicability (or productivity) of a construction is straightforwardly accounted for within the constructional hierarchy. We therefore do not adopt the distinction between constructions and patterns of coining suggested by Kay (2013).

grammaticalisation and lexicalisation or grammatical versus lexical constructionalisation. Some parts of the network are showing signs of what traditionally would be called grammatical constructionalisation (increase in schematicity and productivity). We have found that new subschemas may emerge out of concrete instances (i.e. schema-formation or schematisation) and that already established subschemas may become more productive and more schematic as a result of relaxing collocational constraints and an increase in type frequency (see Ch5, §5.4). At the same time, it is not at all unusual to find that lower levels may display certain conventionalisation or even fossilisation effects that are more akin to lexicalisation processes. Specific instantiations of the construction were found to have increased their token frequencies and developed into conventional, sometimes "fixed" combinations at the lowest lewel of the network. In some cases, the fossilisation of a specific micro-construction may result in a weakening and eventual loss of the overarching lower-level schema (e.g. *zich wild schrikken/ergeren* 'to startle/annoy oneself wild'). Still, there is no causal link between the conventionalisation or fossilisation of a specific instance and the obsolescence of the overarching schema, as we have seen examples of the conventionalisation of new collocations that did not really affect the representation of the superordinate subschema, e.g. the emergence of the conventional collocation *zich suf piekeren* 'to worry oneself drowsy' was more or less concomitant with the creation of the subschema [SUBJ $V_{\text{mental activity}}$ REFL *suf*] and did not prevent the later emergence of the more abstract schema [SUBJ V REFL *suf*]. Clearly, the reshaping of the network is characterised by both schematisation and expansion as well as conventionalisation and, in some cases, loss in specific areas of the network. If one takes a hierarchic network-oriented perspective, these different kinds of developments or constructional changes can be studied in an integrated fashion.

(b) Productivity at different levels of abstraction

One of the main purposes of adopting a constructional network perspective in this thesis was to better understand the interrelatedness between the notions of productivity and schematicity. Concretely, the aim was to investigate how the quantitative and qualitative aspects of productivity play a role in shaping the constructional network and how shifts in productivity may reshape the internal organisation of that network. In general, the constructional network needs to be conceptualised as a dynamic system in which small-scale changes slightly adjust the representation of the hierarchic levels of the constructional network and may over time even cause important reorganisations of its internal structure. Starting out at the level of the micro-construction, we saw how conventional collocations may give rise to productive, partially abstract subschemas. We discussed several examples of individual elements that used to be exclusively intercombined in a sort of fixed expression but which gradually emancipated themselves from that collocation over time: the elements came to attract other collocates and thus

gained some productivity. Of course, some elements may also be prevented (e.g. by high token frequency or statistical preemption, cf. supra) from being used outside of a specific collocation, in which case they remain "stuck" at the micro-construction level. In the early stages of productivity, the new coinages are semantically still highly similar to the "original collocate", i.e. the collocate in the conventional collocation. In the constructional network, the budding productivity is represented as a partially abstract subschema emerging over the micro-construction. Instead of being fully lexically specified, this subschema now has one open slot, but this slot is still semantically constrained. Multiple examples of such collocational expansion and schema-formation were discussed in Chapter 5, §5.3.2.1 and §5.4 (e.g. *suf* 'drowsy', *uit de naad* 'out of the seam', *het vuur uit de sloffen* 'the fire out of the slippers'...). Importantly, the emergence of the superordinate subschema does not necessarily trigger the disappearance of the original micro-construction from the network (as is illustrated by *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers', see the network representation in Ch5, §5.4.3). In many cases, the conventional collocation continues to be used next to the new productive coinages, so it should still have a place in the network structure. Once a subschema is established, it may come to attract more and more types and further diversify its semantic range. If the collocational restrictions are relaxed, i.e. if the productivity is no longer limited to a specific semantic domain, the open slot becomes more schematic and the subschema moves up to a higher level in the hierarchy. As this process is happening in different areas of the network, the network structure becomes increasingly complex as the construction expands its use. We have also discussed some examples of low-level subschemas which decrease in productivity or even cease to be productive. If a subschema no longer attracts new types over an extended period of time, the representation of the subschema may be weakened and eventually lost. In some cases, the obsolescing subschema may leave behind some lexically specified micro-constructions as relics of its former productivity, but the subschema itself is no longer part of the network (see the discussion of *wild* 'wild' in Ch5, §5.3.2.1 and §5.4.4). As was mentioned above, the constructional network straightforwardly accommodates both increase and loss of productivity within one and the same construction. Although we are dealing with an entirely different construction, the findings on the role of productivity in network reorganisations are comparable to what has been observed for the V-*ment* construction by Hilpert (2013). Although Hilpert does not actually propose a visual representation of a network, he does show that the idiosyncrasies of some of the historical developments can more easily be accounted for by (re)considering the development of the V-*ment* construction in terms of subschemas emerging or falling out of use. What is left of this once more elaborate network in present-day English is a number of lexically specified instances, i.e. low-level constructs in the constructional network. These may still spawn analogous formations from time to time, but those have not

triggered the re-emergence of an overarching subschema (a scenario which is admittedly uncommon but not theoretically impossible).

The observation that analogical formations are possible even for models that are not (or no longer) subsumed by a higher-order constructional schema leads Hilpert (2013: 121) to ask "how one can reliably distinguish between new coinages that have been licensed by a constructional schema and those that have been formed through analogy". However, we have already argued in §6.2.2 that there is no qualitative difference between productivity and analogy. The constructional network thus allows for tracking shifts in productivity at different levels of abstraction in the hierarchy: productivity can be based on highly abstract schemas but (sub)constructions may also be extended to new types on the basis of low-level productivity islands or even item-specific extensions. Our data contained several new specific instances that have probably entered the construction through very local, low-level analogies (e.g. the variations on *zich groen en geel ergeren* 'to annoy oneself green and yellow' with infrequent colour terms, the variants of *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers' with other types of footwear, other types of headwear modelled on *een hoedje* 'a little hat', etc.), rather than as "productive" instantiations of a higher schematic pattern. At the same time, we found several examples which can only be explained by the existence of a productive high-level abstract pattern: the productivity at the highest level of abstraction facilitates "random" combinations of lexical items that are compatible with the empty slots at the abstract schema level, even if these specific combinations "overrule" lower-level semantic restrictions or conventional collocations. Still, it appears that most of the instances in the construction could be motivated by productivity at much lower levels in the network. The most important level in our specific construction in terms of productivity was the intermediate level of partially filled subschemas, viz. the intensifier-specific subschemas and verb-specific subschemas. The recent expansion of the construction under investigation was primarily driven by extensions of a number of such highly (token and type) frequent lower-level patterns. In that light, we could suggest that the highest level of schematicity may well be the level that *determines* the overall productivity of a construction (following Barðdal, cf. supra), but it is not the level that necessarily has the highest contribution to – or is primarily "responsible" for – the extensibility of a construction (as is also noted by Barðdal 2008: 98). Although the importance of low-level generalisations (and perhaps even their primacy over highly abstract schemas) has been recognised at least since Langacker (1999), the description of argument structure constructions has generally focused on the highest level of schematicity, with only occasional references to lower levels of abstraction. Recently, however, several studies in (Diachronic) Construction Grammar have started prioritising intermediate levels of abstraction as the main object of study, to the point where higher-level schemas have been claimed not to exist at all (Barðdal et al. 2011, Hilpert 2013, Perek 2015, see Ch2, §2.1.3 for some details). This investigation has also illustrated the value of putting a

greater emphasis on intermediate levels of abstraction both in synchronic studies aimed at comparing productivity of different (sub)constructions and in diachronic studies that track the historical (productivity) development of a construction. At the same time, though, we have illustrated that some specific instances of the construction can only be explained by assuming a highly schematic pattern, so we would not go as far as to argue that higher-level schemas have no role left to play at all (cf. Perek 2015: 114, 141-142).[76]

(c) Potential issues of the constructional hierarchy

To conclude this section, we point out some potential problems or difficulties that come with adopting a Lexicality-Schematicity Hierarchy – or a constructional network approach in general. First, in trying to map out an "accurate" two-dimensional visual representation of the taxonomy that is as true to the data as possible, we ran into some specific problems. As we are building the network from the bottom up, the first natural step is to start out from the actual utterances at the construct level – or at the lexically specified micro-construction level if we abstract away from, e.g., the concrete subject of the clause. As soon as we want to move up to a more abstract intermediate subschema level, we are presented with an important choice: although there may be multiple open slots that play a role in determining the distribution and productivity of the construction (cf. §6.2.2.1 supra), a two-dimensional taxonomy only allows us to abstract over one of these slots at the same time. We could do this according to the hierarchical or nested selection process which assumes that the elements of a clause are generally selected in a fixed order (e.g. first the subject, second the verb, afterwards the objects...) (Zeldes 2012: 125). However, as was argued in §6.2.2.1 this may not be the best way to go for specific constructions. For that reason, we have suggested that it may be useful to build multiple possible representations of the constructional network of one and the same construction, each of which can serve to highlight other generalities in the construction (e.g. focus on different slots or semantic or formal properties at different levels in the hierarchy). Approaching the construction from different, but interacting perspectives in a so-called multirepresentational approach to network structure may provide new insights into the usage peculiarities of some constructions. Even so, it is not easy to decide how, exactly, to build these representations from the bottom up, i.e. how many different levels we should posit (cf. Trousdale 2008a). In Chapter 2, §2.1.3, a four-level distinction between

[76] Perek (2015) finds that the actual usage basis of the conative construction in English is more visible at the level of verb-class-specific constructions. In fact, these local generalisations may better account for the speaker's knowledge of the construction than the higher-level schema. Still, Perek also suggests that the possibility or plausibility of a higher-level schema should not be entirely rejected, as there are some verbs in the conative construction that do not fit in with any of the verb-class-specific constructions (much like was found for some instances of the intensifying fake reflexive resultative construction in this investigation).

constructs, micro-constructions, meso-constructions/subschemas and macro-constructions/schemas was introduced as a heuristic device to describe the taxonomic organisation of a construction. When introducing these levels, Traugott (2008b: 236) mentions that "the hierarchy given here [is not] restricted to these four levels". Our data have also shown that there is no unequivocal universal answer as to what the "ideal" number of levels could be: the appropriate number of levels often needs to be determined on an item-by-item basis and some areas of the network are more densely populated – in the sense of having more different sublevels – than others. In discussing the expansion and reorganisation of the constructional network representation(s), we have tried to illustrate how the frequency and semantic data that we retrieved from our corpus data could help us determine the relative positions and the shifts in the positions of specific micro-constructions and subschemas in the taxonomic hierarchy. Still, we do not know whether our representations really capture the generalisations that are relevant to the language users, as will be discussed in §6.2.3.2 infra.

Second, the changes of the hierarchic structure of the network also bring about reconfigurations of the different kinds of *links* between the nodes in the network, as has been shown by Torrent (2015), among others. In a Lexicality-Schematicity Hierarchy, the primary focus is on the taxonomic inheritance links, i.e. the way in which the different levels within the network are sanctioned or motivated by higher hierarchic levels in the taxonomy. As a first step towards including other types of links in our representations, we introduced interactive links between different representations of the network. These are still a form of taxonomic link, but they cross over between different representations of the same constructional network. They were mainly invoked to illustrate how certain micro-constructions which appeared to be isolated islands or "orphans" (Perek 2015: 142) within one representation, were perfectly motivated by a low-level schema within another possible representation. We also briefly illustrated in what way we could introduce horizontal links into these network representations and what their added value could be. First of all, horizontal links can highlight the paradigmatic relationship between nodes at the same level of abstraction and thus capture certain low-level similarities that were "ignored" due to specific choices made in building the taxonomy from the bottom-up. In a way, these horizontal links in one representation were said to correspond to taxonomic inheritance links in another representation of the network which is based on different generalisations. Horizontal links were also introduced to highlight how some micro-constructions were more tightly related to each other than others, even if they were all subsumed by the same overarching subschema. From a diachronic point of view, these "analogical links" allow us to distinguish between micro-constructions that have most probably entered the network as a direct (productive) instantiation of a higher-level subschema and micro-constructions that seem rather more likely to have entered the network through low-level local analogies (cf. the different types of extension strategies, supra). These different historical origins are not always evident in the taxonomic

hierarchy and the inheritance links, because any micro-construction immediately becomes taxonomically licensed by an overarching schema as soon as it joins the network. As has been shown by Torrent (2015), a new node starts building (inheritance or other) relations with the already existing nodes in the network once it starts "participating" in that constructional network. As any representation of a constructional network is a snapshot at one specific moment in time, the taxonomic links primarily reflect the motivations or generalisations language users establish between those constructions at that time, regardless of their historical origins.

As discussed in Chapter 4, §4.4.1, our proposal for horizontal links is inspired by the existing literature on horizontal links, but it is different in several respects. First of all, the fact that horizontal links may capture a paradigmatic relationship between subschemas is found in Van de Velde (2014), who introduces horizontal links between constructions that form a kind of paradigm, i.e. "a set of alternating forms with related meaning differences" (2014: 149).[77] However, the paradigmatic relationships he discusses are syntactic in nature, in the sense that the alternating syntactic forms have different meanings/functions. For example, he discusses how the position of the verb in the main clause is linked to different kinds of clause types (e.g. V2 in declaratives, V1 for questions, Vfinal for subordinates). During language change, such syntactic paradigms may come under pressure (e.g. in English, all clause types have become V2), but in some cases the semantic or functional differences between horizontally related nodes may survive. The paradigmatic relationship between the subschemas in the intensifying fake reflexive resultative construction should be interpreted more in terms of "interchangeability". The intensifier-specific subschemas that are horizontally linked are in fact perfectly interchangeable without any important change in meaning (e.g. *zich rot/dood/kapot/suf werken* all mean 'to work very hard') and the verb-specific subschemas are interchangeable with only a change in verb meaning (e.g. *zich rot lachen/werken/zoeken* means 'to laugh/work/search intensely'). In other previous work, horizontal links have mainly been invoked in order to elucidate how speakers may recognise formal or semantic similarities between different constructions that are not taxonomically related and how analogisation on the basis of those cross-constructional similarities may lead to constructional change. De Smet & Fischer (2017) and De Smet et al. (2018) suggest that changes within a specific construction should be interpreted against the background of a broader constructional network. De Smet & Fischer (2017) show, for instance, that certain changes may be facilititated if there are other ("supporting") constructions available on which the innovation can be modelled, and, conversely, a change may be inhibited if it is

[77] A different use of horizontal links is found in Taylor (2004) , who argues that elements enter into horizontal, syntagmatic relations with the larger structure of which they are part. The horizontal links that were introduced in this thesis, however, capture the *paradigmatic* rather than the syntagmatic relationship between the subschemas.

not supported by other constructions (see also Abbot-Smith & Behrens 2006). In their discussion on the possible outcomes of linguistic competition, De Smet et al. (2018) suggest that attraction may follow from horizontal links (i.e. analogy) between functionally similar forms; differentiation, on the other hand, may be interpreted as overlapping forms becoming less similar because they are aligned with and taxonomically motivated by distinct subnetworks. In some cases, constructions which share a certain (superficial) formal and semantic resemblance only *occasionally* "contaminate" each other (i.e. in some, but not all realisations of the construction) (Pijpops & Van de Velde 2016). Over time, such contamination or cross-constructional analogy may result in actual change or even form the basis of so-called multiple source constructions, which are the result of the merger or conflation of a number of patterns (either at a macro-level or at a micro-level) with different lineages, which were apparently confounded or at least perceived to be highly similar (see, e.g., Israel 1996, Verhagen 2003a, b, and the papers in De Smet et al. 2015 for some examples). It would be interesting to explore in future research how our horizontal links between subschemas *within* the Lexicality-Schematicity Hierarchy of one construction can be extended to links with (sub)constructions in other constructional hierarchies, and how some of the nodes in the current network representations are possibly influenced by such cross-constructional links.

In the current paragraph, we have described the constructional network as a theoretical construct that can account for a substantial amount of the variation and change attested in constructions. In other studies, several of the concepts that we have touched upon here have been related to certain cognitive processes, although the distinction between what counts as a "linguistic" phenomenon and what counts as a "cognitive" effect is often left implicit. The process of schematisation, through which higher-order abstractions emerge over specific instances of a construction, rests on the idea that language users recognise similarities between these specific instances and abstract away from the shared elements in the form of a generalisation (Bybee 2006, 2010). However, while the linguist can posit certain similarities between instances of the same construction or between different constructions at a high level of granularity, it is by no means certain that the language user recognises these similarities as well (Perek 2015). At the micro-construction level, we have seen that some instances of a construction are more frequent and conventional than others. As both the degree of entrenchment and the "conventional" status of specific linguistic units appear to hinge on token frequency, a certain shortcut between conventionalisation and entrenchment has been suggested (Schmid 2017: 15-16). Given the link between type frequency and entrenchment at the schema level, then, entrenchment has been tied up with productivity (Bybee 1995, Clausner & Croft 1997, Barðdal 2008, Langacker 2008). Another phenomenon that invites certain cognitive implications is analogy. Although we have thought of analogical extensions as a primarily linguistic phenomenon, analogisation

implicitly assumes a form of analogical thinking and the recognition of certain similarities between different forms on the part of the language user (cf. Traugott & Trousdale 2013, who distinguish between the (cognitive) process of analogical thinking, i.e. the matching of formal and semantic aspects, and analogisation as a linguistic mechanism). As the natural links that are often assumed between linguistic phenomena attested in corpora and cognitive processes in the minds of the speaker have recently met with some scepticism, we further address this issue in the next paragraph.

6.2.3.2 The cognitive status of the constructional hierarchy

In Chapter 2, it was mentioned that Construction Grammar is an umbrella term that covers multiple construction-based approaches that share certain tenets but differ in other (subtle or more important) respects. Most construction grammars share a certain cognitive or mentalistic outlook on language in that they aim to account for the language user's knowledge of language. In usage-based constructionist approaches, it is assumed that this knowledge of language is formed by and constantly reformed through language use. In order to study language use and gain access to the cognitive representation of language, usage-based studies often rely on corpus data. If Diachronic Construction Grammar is to be a diachronic application of Construction Grammar, i.e. a framework for studying language change from a constructional perspective, it should also be able to account for how the *knowledge* of language has changed over time. Nevertheless, it appears that the main concern of studies taking a diachronic constructional approach is to document the changes in form, meaning or frequency of specific constructions, without really explicitly addressing the cognitive implications of any of the observed changes – in fact "any claims about the linguistic knowledge of earlier generations of speakers stands on rather shaky ground, given the limited representativeness of historical corpora" (Hilpert 2018: 23). Whether a more explicit cognitive commitment or references to psychological reality should be a priority in the research field of Diachronic Construction Grammar is not a question that is easily answered (see also von Mengden & Coussé 2014). Hilpert merely suggests that, *if* the researcher wishes to establish a connection between language use and cognition, this goal should be explicitly signalled. Let us therefore carefully consider to what extent the changes that were discussed in this thesis may also reveal something about the way in which the language user's knowledge or cognitive representation of the construction has developed over the past two centuries.

Our study has first and foremost treated the constructional network as a taxonomic representation of the *use* of one specific construction, as can be gleaned from corpus data. In describing the hierarchic structure and internal reorganisations of the network, we have primarily referred to linguistic phenomena such as quantitative and qualitative productivity, schematicity/schematisation or conventionality/conventionalisation (cf. the previous paragraph). In Chapter 4 we have emphasised that it is crucial to distinguish

between the cognitive representation of the network, which should include all relevant aspects of the use of a specific construction, and the linguist's (visual) representation of the constructional network. However, that does not mean that the two conceptualisations of the constructional network are two ontologically different entities. While the taxonomic representation of the network may not be a perfect one-to-one mapping of the cognitive configuration of the network, we have argued that the aim is (or should be) to give at least *some* indication of the cognitive representation of that particular construction in the minds of language users, with the important caveat that this representation is a simplification which abstracts away from differences between individual language users. In that regard, the different levels of representation in the Lexicality-Schematicity Hierarchy (cf. supra) to some extent reflect that there is a difference between the cognitive representations of "fixed expressions" (or conventional collocations) at the lowest level, intermediate patterns that can only be instantiated by a limited number of types and higher-level patterns that can host a wide range of types. In building the taxonomic network representations we took a usage-based perspective in that we were careful to only posit certain higher-level abstractions if there was sufficient evidence in the corpus that such generalisations may have been made by at least some speakers. That is, although the nodes at different levels of abstraction are primarily intended as theoretical constructs, they should reflect plausible generalisations in the minds of the speakers. Of course, this raises several questions. First of all, we could ask ourselves what counts as "sufficient" evidence: how many tokens and types does a language user need to have encountered in order for him/her to notice certain similarities between these specific instances which can then be captured in a higher-order generalisation? The question implies that there is a critical number of types needed to enable the creation of a schema, but the problem with that is that it is not fully clear whether the formation of a new schema should be seen as an instantaneous event or as a gradual process. That is, while we can track the emergence of new subschemas over an extended period of time in a large-scale diachronic corpus, it seems impossible to determine the exact moment at which a set of micro-constructions has triggered the formation of a higher-order subschema. Nonetheless, the question is important in that it brings attention to the fact that the linguist needs to be aware that any configuration of schemas and subschemas he/she has constructed on the basis of corpus data is open to discussion. It is possible that the linguist has posited certain intermediate subschemas or structural links between the nodes in the network that are not actually relevant to the language user. This brings us to the second question: what kind of similarities (e.g. formal or semantic) are perceived by the speaker and which kind of generalisations does he/she make? As a first step towards accounting for the fact that speakers may actually perceive of different types of similarities or arrive at several generalisations simultaneously, which entails slightly different configurations of the network, we have suggested multiple interacting network representations of one and the same construction and we have added

horizontal links to capture certain potentially relevant similarities that are not captured by the taxonomic structure (cf. supra). Even so, our two-dimensional visualisations are not fully able to capture the necessary complexity and dynamism of the constructional network.

In addition, we have also tentatively provided an indication of the strength of the cognitive representation of different levels of abstraction in the hierarchy. Entrenchment is often mentioned as an effect of high discourse frequency in that a linguistic unit that is often repeated may come to be cognitively routinised. While this is the most common use of entrenchment, it was mentioned that (partially) abstract patterns may also be entrenched, in which case the type frequency – rather than token frequency – is an indicator of the degree of entrenchment. In our investigation, the two types of entrenchment were incorporated as entrenchment of lexically specified micro-constructions (i.e. token entrenchment) and entrenchment of partially schematic patterns (i.e. type entrenchment). It needs to be pointed out that, as our network representations are super- and supra-individual abstractions, we are referring to entrenchment at the level of the *average* language user; the exact degree of entrenchment is of course different for each individual language user, depending on his/her linguistic experience. Although the direct inference of entrenchment from corpus frequencies has been questioned or reconsidered recently (Schmid 2010, Blumenthal-Dramé 2012), there is some logic to it. The fact that a particular lexically specified unit is highly frequent – or even more frequent than one would expect, if one uses collostruction strength rather than absolute token frequencies as an indicator of entrenchment (Stefanowitsch & Flach 2017) – naturally invites the assumption that it must have some kind of unit-status in the minds of the speaker. Conversely, if a certain partially abstract pattern occurs with a wide range of types, the language user seems to have internalised some kind of open-ended schema. While the idea that language units can be stored both analytically and holistically is well-accepted since Langacker (1987), it is still sometimes maintained that entrenched units, even though they formally conform with a more abstract schema, are by necessity not assembled from scratch, and vice versa (Dąbrowska 2004: 20, Taylor 2004: 15). Clausner & Croft (1997), for example, have suggested that either the tokens are more entrenched than the overarching schema or, conversely, the schema is more entrenched than the individual instances. However, we have illustrated that entrenchment at the token level and entrenchment at the type level are not necessarily incompatible, by discussing several examples of entrenched micro-constructions that were also licensed by an entrenched subschema (e.g. *zich suf piekeren* 'to worry oneself drowsy', *zich uit de naad werken* 'to work oneself out of the seam', see the network representations in Ch4, §4.4 and Ch5, §5.4). Indeed, it is not because a specific combination is extremely frequent, and thus perhaps stored and entrenched as a chunk, that a language user is unable to perceive the internal structure of this combination. In the synchronic part of our investigation, the fact that certain well-entrenched collocations are not fully fixed and

often display a small amount of lexical variation suggests that language users do recognise the internal structure of these collocations or are to some extent aware that they are in fact instantiations of a higher-level abstract pattern (cf. supra). The diachronic analysis then showed that this combination of well-entrenched micro-constructions and overarching entrenched subschemas is often a natural result of a construction's historical development. Several of the entrenched subschemas in our present-day network representations were found to have developed from such entrenched micro-constructions (cf. the discussion on how small-scale analogical extensions may give rise to productivity, supra).

This brings us to the topic of how we should interpret entrenchment from a diachronic point of view, i.e. how the degree of entrenchment of specific units changes over time and what the implications of such changes may be. A number of specific phenomena in language change have been related to the notion of entrenchment. In his discussion of the role of entrenchment in language change, De Smet (2017) finds that entrenchment can sometimes serve as a conservative force, like when a well-entrenched unit resists analogical levelling or regularisation, but that it can also be at the basis of linguistic innovations, for instance when the perceived similarities between different constructions lead to constructional change (cf. supra). He also concludes that our understanding of entrenchment and the way in which it interacts with other mechanisms is still incomplete. In our investigation as well, it appears that some units are so entrenched that they have remained unchanged for an extended period of time (conservation), while other entrenched units have given rise to new creative variations and subschemas within the network (innovation, cf. the ambiguous effects of token frequency in §6.2.2). How the degree of entrenchment itself changes over time is another question. At first blush, the diachronic dimension of entrenchment seems rather straightforward: if entrenchment is an effect of high frequency, changes in frequency should affect the degree of entrenchment of different linguistic units. At the token level, repeated use contributes to a higher degree of entrenchment, whereas extended periods of disuse negatively influence the degree of entrenchment (Langacker 1987: 59). At the schema level, new types may serve as additional proof of the existence of an abstract schema, thus strengthening its cognitive representation. Conversely, if a schema comes to be attested with fewer and fewer types over time, the representation of the schema becomes weaker, sometimes causing the schema to fall out of use entirely. In a constructional network, in which patterns are taxonomically organised from highly abstract to highly specified, it is not always easy to know which level is activated by specific instances of the construction. In §5.4.5 we discussed the upward-strengthening-hypothesis by Hilpert (2015a), which makes certain claims about the strengthening at different levels of abstraction. He proposes that not all instances of a construction necessarily contribute to the strengthening of the highest level schema. For one, there may be a prominent intermediate subschema that prevents further upward

strengthening. In his case study on noun-participle compounding, Hilpert (2015: 138) suggests that the encounter of a word like *Stratocaster-based* will strengthen the [N-*based*] subschema, but it might not reach the [N-participle] schema. Second, upward strengthening may also be hampered by high frequency. If a specific linguistic unit is extremely frequent, another use of that linguistic unit will further strengthen the cognitive entrenchment of that unit but not (necessarily) of the more abstract schema. An example of such an entrenched noun-participle compound is *home-made*, which is much less likely to strengthen the [N-participle] subschema than an infrequent form like *oxen-yoked* (Hilpert 2015: 138). Diessel (2004: 30) even argues that, not only does a new instance of a highly entrenched type not strengthen the more abstract schema, "very high token frequency can *weaken* the connection of a type to a constructional schema" (emphasis added). In other words, the assumption appears to be that certain instances do not strengthen more abstract schemas, because they are not (primarily) *recognised* as instances of those abstract schemas. However, we have provided both synchronic and diachronic indications that language users do have access to the internal structure, even in the case of entrenched micro-constructions, and that they can project upwards beyond entrenched intermediate subschema levels. For instance, the fact that language users may (deliberately) overrule certain low-level semantic restrictions suggests that they are aware of the existence of a higher-order abstract schema beyond the (entrenched) intermediate, semantically constrained subschema. Over time, such creative uses may cause the semantic restriction to be relaxed. In addition, we saw that certain "fixed" collocations may display some amount of lexical variation and that this kind of variation may even give rise to a new low-level subschema. This was said to indicate that language users do process these collocations as proper instances of the [SUBJ V REFL INT] schema, rather than as independently stored chunks. While the fact that they have *access* to the internal structure should not be taken to mean that the internal structure is necessarily always activated, it does invite some caution when making claims about which levels are and are not strengthened by specific instances of a construction.

An interesting approach to entrenchment from a Construction Grammar perspective comes from Hilpert & Diessel (2017), who focus on the entrenchment of the links between different nodes in the constructional network, rather than the entrenchment of the nodes themselves. With respect to inheritance links, they suggest that more prototypical specific instances of an abstract pattern (e.g. instances of the DOC with the verb *give*) have stronger, more entrenched links to the higher-order schema and are more easily recognised as instances of that schema than less prototypical instances are. Furthermore, constructions that have highly entrenched subpart links are more transparent and more easily analysed into their component parts than constructions with weaker subpart

links.[78] An example of the latter is *let alone*, which is not parsed as an instance of a higher-order imperative verb schema but as a standalone item. In addition, they treat the entrenchment of collocations or collostructions in a slightly different way as well. Rather than saying that items with a high frequency of co-occurrence are entrenched as a unit, it is the *link* between the items that is entrenched. If the link between two lexical items or between a lexical item and a construction is not entrenched at all, the combination may be highly unexpected and in some cases even judged as unacceptable (so-called negative entrenchment, Stefanowitsch 2008). In a way, the covarying collexeme analysis (and other collostructional analyses), which we used to uncover preferred and dispreferred collocational patterns, could then also be reinterpreted as measuring the entrenchment of the link between two lexical items. The higher the entrenchment of the link between the items, the higher the collostruction strength; in case of negative entrenchment, the collostruction strength is also negative and we get repelled collocations. The discussion in Hilpert & Diessel (2017) is in line with a recent suggestion that the connections or links between constructions are perhaps more interesting and better suited as a model for the cognitive organisation of constructions than the nodes themselves. In fact, the whole idea of a hierarchic network structure is put under review by Schmid (2017):

> They [usage-based models] claim that these constructions and schemas are related to each other in a massive associative memory network organized mainly in terms of hierarchical relations. The present proposal diverges from this idea in two important ways: First, it rejects the distinction between constructions serving as nodes in the network and relations between nodes and instead assumes that linguistic knowledge is available in one format only, namely, associations. These associations come in four types: symbolic, syntagmatic, paradigmatic, and pragmatic. (Schmid 2017: 25)

The main problem, according to Schmid, is that the network is often conceptualised as a static repository of nodes, which are posited as abstractions over specific instances. Although these nodes are in themselves complex units, this internal structure is not obvious in the network organisation. It is true that, in our taxonomic representations of the constructional network, the representation and hierarchical position of certain subschemas remained largely identical over time, and we did not really elaborate on how certain subschemas could differ from one another with regard to their internal structure and there are certain types of information that are relevant for the actual usage of the construction but which we did not directly account for in our network representations, often due to limitations imposed by the chosen visualisation method. In order to better

[78] Subpart links are meronomic links which indicate that one construction, although existing independently, is a proper subpart of another construction (Goldberg 1995: 78).

account for the dynamism of the constructional network, the internal structure of which is constantly in flux and may over time undergo important reorganisations, greater emphasis should be placed on the connections or links between the nodes. There have already been some attempts to reconcile a more connectionist view on constructions with the traditional view of networks. The case study by Torrent (2015), which was discussed in Chapter 2, §2.1.3, deals with both the emergence of new nodes within the network and the reconfiguration of the links between the nodes, in order to show how the internal configuration of the network accommodates to the emergence of new nodes. Crucially, Hilpert (2018: 33) points out that the two views are not incompatible: while they may focus on different aspects of constructions and constructional change, they essentially capture the same insights. In that sense, reducing all constructions to a set of associations and completely abandoning the idea of nodes, as Schmid (2017) suggests, may be somewhat too drastic – especially since Schmid's proposal is still underdeveloped, both from a theoretical and an empirical point of view. The previous subsection has illustrated how several constructional changes can be couched in terms of shifts in the network, suggesting that a taxonomic network structure does have some inherent appeal, at least as a theoretical construct. In this subsection, we have further suggested that the network representations may have some cognitive reality, but there is of course still a lot of work to be done before we can truly bridge the gap between corpus data and cognition.

6.3 Directions for further research

The previous section discussed how the results presented in this thesis may contribute to our knowledge on the mechanisms and factors involved in constructional productivity and network (re)organisations within the larger framework of Diachronic Construction Grammar. At the same time, we also laid bare some potential shortcomings of this investigation, which could be addressed in further research.

First of all, this investigation has primarily focused on the language-internal factors that may play a role in productivity and the constructional network (re)organisations. However, there are still some uncertainties with respect to the exact implications of some of these factors, especially with regard to the ambiguous role of token frequency. We argued that high token frequency may either contribute to or detract from productivity, but that it was generally difficult to predict which of the two outcomes we would get. Of course, it is possible that there are other factors interacting with or even confounding the effect of high token frequency that were not considered in the current investigation. In general, the methodological toolkit that was developed to study the different facets of productivity proved to be quite successful in highlighting the role of different types of

frequency information and semantic aspects in productivity, although the concrete measures still need further fleshing out. For example, the role of semantics in constructional productivity could be better accounted for if we could operationalise or find a better empirical foundation for the notion of semantic coherence or variability. It was suggested that methods from distributional semantics, such as the Word Space models used by Perek (2016a, 2016b) may make it possible to provide a more fine-grained picture of the semantic development of a construction, provided that the corpus allows for such an approach (i.e. the corpus needs to be enriched with PoS-tagging and OCR-mistakes should be reduced to a minimum). Based on the coarse semantic categories that were distinguished in this investigation, we were already able to map out the broad semantic range (and expansion) of the construction and capture some of the relevant collocational restrictions at lower levels, but it would be interesting to see whether new insights can be gained from submitting the construction and its subschemas to a more detailed semantic analysis.

In addition, we could expand on the current investigation with a more elaborate variationist dimension in order to shed some light on the role of extra-linguistic factors on the productivity of constructions. While we did explore the impact of national variation on the productivity and constructional network organisation in our synchronic study, the diachronic part of the investigation only focused on Netherlandic Dutch. A first obvious step would be to construct a diachronic data set for Belgian Dutch as well, which is an enterprise that would first require the compilation of a large diachronic corpus of Belgian Dutch. In light of the different standardisation processes in Belgian and Netherlandic Dutch, it would be interesting to see whether the network structures in the two national varieties show evidence of convergence or divergence. There were some indications that the Belgian Dutch variant of the construction is more "conservative" in the sense that some intensifiers which have (virtually) disappeared from Netherlandic Dutch are still present in Belgian Dutch. It might be the case that some intensifiers were first introduced in Netherlandic Dutch before being adopted by Belgian speakers and that, in the same vein, they were quicker to fall out of use in Netherlandic Dutch than in Belgian Dutch – but this does not explain why Belgian Dutch also has a slightly higher proportion of (creative) hapax intensifiers. At any rate, certain reorganisational shifts may be happening at different times or at a different pace in the national variants of the constructional network. Second, as our investigation is based on data from one specific genre (which in itself has undergone some changes that may have influenced the results to some extent, cf. §5.5.2), it would be useful to compare the use and productivity of the construction in different genres. In the tradition of morphological productivity, some word-formation patterns have been found to be more productive in one register than in another (Plag et al. 1999). Given the lack of historical data for most other genres, we currently do not have the necessary tools for comparing different genres from a diachronic perspective, but we can at least consider the impact of register on synchronic

productivity. Although the majority of the data was accounted for by conventional combinations of highly frequent verbs and intensifiers, the journalistic data also contained many creative and (deliberately) unconventional instances of the construction, suggesting that the construction is highly productive in journalese. Judging by the sporadic examples from Twitter or the list of (Internet) intensifiers listed on the blog *Pelikanenschurft* (see Appendix III-1), however, language users appear to be even more resourceful in informal registers, which could indicate that the construction at the highest level of schematicity is more productive in informal language. A natural further step is to include even more extra-linguistic variables in the (synchronic) study of constructional productivity. Two of the more interesting factors to take into account are gender and age of the speaker. It was mentioned in passing that women are often considered to be linguistic innovators and have been reported to come up with new intensifiers more readily than male speakers. It would be interesting to see whether women indeed have a larger repertoire of intensifiers or are more willing to break with convention in this construction as well. The use of intensifiers has also been associated with younger generations, who are generally more sensitive to linguistic fashions and have been shown to use specific intensifiers as identity markers that are not shared by the broader linguistic community (see, e.g., Stenström 1999, Ito & Tagliamonte 2003, De Clerck & Colleman 2013, Pertejo & Martínez 2014). It is not unlikely that the use of the intensifying fake reflexive resultative construction is subject to similar age effects. If we could find a systematic way to account for such lectal variation in the study of constructional productivity, we would be able to study productivity at a more inter-individual level and, accordingly, add a variationist dimension to the constructional network.

As our understanding of the factors that play a role in (re)shaping the constructional network increases, it will also become more important to refine the (taxonomic) network representations that were presented in this investigation. Recent years have seen a growing interest in finding a better computational basis for Construction Grammar by incorporating some of its main tenets into computational models. Given dynamic, interactive constructional networks, in which we are less hampered by certain visual limitations, we may be able to encompass more details on the use of a construction in our networks. This would also allow us to further look into the different types of links between the nodes in the network and how these links may be reconfigured over time. As was argued in the previous paragraph, there are important unresolved questions with respect to the extent to which corpus data can inform us about the cognitive representation of language in individual speakers. A more connectionist approach to constructional networks may be more suited to account for the cognitive implications of the constructional network.

Bibliography

Aaron, Jessi E. 2016. The Road Already Traveled: Constructional Analogy in Lexico-Syntactic Change. *Studies in Language* 40, 26-62.

Aarts, Bas. 1989. Verb-Preposition Constructions and Small Clauses in English. *Journal of Linguistics* 25, 277-290.

Aarts, Bas. 1995. Secondary Predicates in English. In: Bas Aarts & C. F. Meyer (eds.), *The Verb in Contemporary English*, 75-101. Cambridge: Cambridge University Press.

Abbot-Smith, Kirsten & Heike Behrens. 2006. How Known Constructions Influence the Acquisition of Other Constructions: The German Passive and Future Constructions. *Cognitive Science* 30 (6), 995-1026.

Adams, Francis M. & Charles E. Osgood. 1973. A Cross-Cultural Study of the Affective Meanings of Color. *Journal of Cross-Cultural Psychology* 4 (2), 135-156.

Agresti, Alan. 2002. *Categorical Data Analysis (2nd Ed.)*. New York: Wiley.

Allerton, David J. 1987. English Intensifiers and Their Idiosyncrasies. In: Ross Steele & Terry Threadgold (eds.), *Language Topics: Essays in Honour of Michael Halliday*, 15-32. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Anshen, Frank & Mark Aronoff. 1999. Using Dictionairies to Study the Mental Lexicon. *Brain and Language* 68, 16-26.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark. 1983. A Decade of Morphology and Word Formation. *Annual Review of Anthropology* 12, 355-375.

Audring, Jenny & Geert Booij. 2016. Cooperation and Coercion. *Linguistics* 54 (4), 617-637.

Baayen, Harald R. 1989. *A Corpus-Based Approach to Morphological Productivity. Statistical Analysis and Psycholinguistic Interpretation*. Unpublished PhD. dissertation, Vrije Universiteit Amsterdam.

Baayen, Harald R. 1990. Corpusgebaseerd Onderzoek Naar Morfologische Produktiviteit. *Spektator* 19, 213-233.

Baayen, Harald R. 1992. Quantitative Aspects of Morphological Productivity. In: Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology*, 109-149. Dordrecht: Kluwer Academic Publishers.

Baayen, Harald R. 1993. On Frequency, Transparency and Productivity. In: Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology*, 181-208. Dordrecht: Kluwer Academic Publishers.

Baayen, Harald R. 2009. Corpus Linguistics in Morphology: Morphological Productivity. In: Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics. An International Handbook.*, 900-919. Berlin/New York: Mouton De Gruyter.

Baayen, Harald R. & Rochelle Lieber. 1991. Productivity and English Derivation: A Corpus-Based Study. *Linguistics* 29 (5), 801-844.
Barðdal, Jóhanna. 2006. Construction-Specific Properties of Syntactic Subjects in Icelandic and German. *Cognitive Linguistics* 17 (1), 39-106.
Barðdal, Jóhanna. 2007. The Semantic and Lexical Range of the Ditransitive Construction in the History of (North) Germanic. *Functions of Language* 14, 9-30.
Barðdal, Jóhanna. 2008. *Productivity Evidence from Case and Argument Structure in Icelandic*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Barðdal, Jóhanna. 2009. The Development of Case in Germanic. In: Jóhanna Barðdal & Shobhana L. Chelliah (eds.), *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, vol. 108, 123-159. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Barðdal, Jóhanna. 2011. The Rise of Dative Substitution in the History of Iceland: A Diachronic Construction Grammar Account. *Lingua* 121 (1), 60-79.
Barðdal, Jóhanna. 2013. Construction-Based Historical-Comparative Reconstruction. In: Graeme Trousdale & Thomas Hoffmann (eds.), *Oxford Handbook of Construction Grammar*, 438-457. Oxford: Oxford University Press.
Barðdal, Jóhanna & Spike Gildea. 2015. Diachronic Construction Grammar: Epistemological Context, Basic Assumptions and Historical Implications. In: Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic Construction Grammar*, 1-49. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Barðdal, Jóhanna, Kristian E. Kristoffersen & Andreas Sveen. 2011. West Scandinavian Ditransitives as a Family of Constructions: With a Special Attention to the Norwegian V-Refl-Np Construction. *Linguistics* 49 (1), 53-104.
Barðdal, Jóhanna et al. 2015. *Diachronic Construction Grammar*.
Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
Bauer, Laurie & Winifred Bauer. 2002. Adjective Boosters in the English of Young New Zealanders. *Journal of English Linguistics* 30 (3), 244-257.
Bergs, Alexander & Gabriele Diewald. 2008. *Constructions and Language Change*. Berlin/New York: Mouton De Gruyter.
Biber, Douglas. 2004. Historical Patterns for the Grammatical Marking of Stance: A Cross-Register Comparison. *Journal of Historical Pragmatics* 5 (1), 107-136.
Biber, Douglas & Edward Finegan. 2001. Diachronic Relations among Speech-Based and Written Registers in English. In: Susan Conrad & Douglas Biber (eds.), *Variation in English: Multi-Dimensional Studies*, 66-83. London/New York: Routledge.
Blanco-Suárez, Zeltia. 2010. On the Origin and Grammaticalisation of the Intensifier *Deadly* in English. Paper presented at Langwidge Sandwidge, University of Manchester.
Blumenthal-Dramé, Alice. 2012. *Entrenchment in Usage-Based Theories. What Corpus Data Do and Do Not Reveal About the Mind* (Topics in English Linguistics). Berlin/New York: Mouton De Gruyter.
Boas, Hans C. 2003. *A Constructional Approach to Resultatives*. Stanford, CA: CSLI Publications.
Boas, Hans C. 2005. Determining the Productivity of Resultatives: A Reply to Goldberg and Jackendoff. *Language* 81 (2), 448-464.
Boas, Hans C. 2008. Resolving Form-Meaning Discrepancies. In: Jaako Leino (ed.), *Constructional Reorganization*, 11-36. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Boas, Hans C. 2011. Coercion and Leaking Argument Structures in Construction Grammar. *Linguistics* 49 (6), 1271-1303.
Bolinger, Dwight L. 1972. *Degree Words*. Den Haag/Paris: Mouton.

Boogaart, Ronny, Timothy Colleman & Gijsbert Rutten. 2014. *Extending the Scope of Construction Grammar*. Berlin/New York: Mouton De Gruyter.
Booij, Geert. 1977. *Dutch Morphology. A Study of Word Formation in Generative Grammar*. Lisse: The Peter de Ridder Press.
Booij, Geert. 2009. Productiviteit, Creativiteit En Écriture Artiste. In: Ronny Boogaart, Josien Lalleman, Marijke Mooijaart & Marijke Van der Wal (eds.), *Woorden Wisselen*, 27-36. Leiden: Stichting Neerlandistiek Leiden.
Booij, Geert. 2010a. *Construction Morphology*. Oxford: Oxford University Press.
Booij, Geert. 2010b. Construction Morphology. *Language and Linguistics Compass* 4 (7), 543-555.
Börjars, Kersti, Nigel Vincent & George Walkden. 2015. On Construction a Theory of Grammatical Change. *Transactions of the Philological Society* 113 (3), 363-382.
Bowers, John. 1997. A Binary Branching Analysis of Resultatives. Paper presented at The 1997 Texas Linguistics Society Conference: The Syntax and Semantics of Predication, Austin: University of Texas.
Boyd, Jeremy K. & Adele E. Goldberg. 2011. Learning What Not to Say: The Role of Statistical Preemption and Categorization in *a*-Adjective Production. *Language* 87 (1), 55-83.
Braine, Martin D. S. & Patricia J. Brooks. 1995. Verb Argument Structure and the Problem of Avoiding an Overgeneral Grammar. In: Michael Tomasello & William E. Merriman (eds.), *Beyond Names for Things: Young Children's Acquisition of Verbs*, 353-376. Hillsdale, NJ: Lawrence Erlbaum.
Broccias, Cristiano. 2004. The Cognitive Basis of Adjectival and Adverbial Resultative Constructions. *Annual Review of Cognitive Linguistics* 2, 103-126.
Broekhuis, Hans. 2013. *Syntax of Dutch. Adpositons and Adpositional Phrases*. Amsterdam: Amsterdam University Press.
Broekhuis, Hans, Norbert Corver & Riet Vos. 2015. *Syntax of Dutch. Verbs and Verb Phrases Vol 1.* Amsterdam: Amsterdam University Press.
Brooks, Patricia J. & Michael Tomasello. 1999. How Children Constrain Their Argument Structure Constructions. *Language* 75 (4), 720-738.
Bürgisser, Max. 1983. *Untersuchungen Zur Wortbildung Im Althochdeutschen Und Altniederdeutschen: Form Und Funktion Von Denominalen Ableitungen in Der Benediktinerregel, Im Tatian Und Im Heiland*. Bern/New York: P. Lang.
Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Bybee, Joan. 1995. Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10 (5), 425-455.
Bybee, Joan. 2003. Mechanisms of Change in Grammaticization: The Role of Frequency. In: Brian D. Joseph & Richard D. Janda (eds.), *The Handbook of Historical Linguistics*. Oxford: Blackwell.
Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82 (4), 711-733.
Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
Bybee, Joan & David Eddington. 2006. A Usage-Based Approach to Spanish Verbs of 'Becoming'. *Language* 82 (2), 323-355.
Bybee, Joan & Sandra Thompson. 1997. Three Frequency Effects in Syntax. *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Pragmatics and Grammatical Structure* 23, 378-388.
Campbell, Lyle. 2001. What's Wrong with Grammaticalization? *Language Sciences* 23, 113-161.
Cappelle, Bert. 2011a. Erop Los Intensifiëren. *Over Taal* 50 (2), 38-39.
Cappelle, Bert. 2011b. Er Verder Op Los Intensifiëren. *Over Taal* 50 (3), 66-67.

Cappelle, Bert. 2014. Conventional Combinations in Pockets of Productivity: English Resultatives and Dutch Ditransitives Expressing Excess. In: Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*, 251-282. Berlin/New York: Mouton De Gruyter.
Charteris-Black, Jonathan. 2004. Metaphor in Sports Reporting. In: Jonathan Charteris-Black (ed.), *Corpus Approaches to Critical Metaphor Analysis*, 113-134. London: Palgrave Macmillan UK.
Charteris-Black, Jonathan & Andreas Musolff. 2003. 'Battered Hero' or 'Innocent Victim'? A Comparative Study of Metaphors for Euro Trading in British and German Financial Reporting. *English for Specific Purposes* 22 (2), 153-176.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Church, Kenneth Ward & Patrick Hanks. 1990. Word Association Norms, Mutual Information and Lexicography. *Computational Linguistics* 16 (1),
Claridge, Claudia. 2011. *Hyperbole in English: A Corpus-Based Study of Exaggeration*. Cambridge: Cambridge University Press.
Clarke, Tom & Alan Costall. 2008. The Emotional Connotations of Color: A Qualitative Investigation. *COLOR Research and application* 33 (5), 406-410.
Claudi, Ulrike. 2006. Intensifiers of Adjectives in German. *Sprachtypologie und Universalienforschung* 59 (4), 350-369.
Clausner, Timothy C. & William Croft. 1997. Productivity and Schematicity in Metaphors. *Cognitive Science* 21 (3), 247-282.
Clement, Marja & Elvira Glaser. 2014. Hoe Krijg Je Dit Gepubliceerd? 'Krijgen'-Constructies in Het Nederlands. In: Freek Van de Velde, Hans Smessaert, Frank Van Eynde & Sara Verbrugge (eds.), *Patroon En Argument: Een Dubbelfeestbundel Bij Het Emeritaat Van William Van Belle En Joop Van Der Horst*. Leuven: Universitaire Pers Leuven.
Colleman, Timothy. 2010. Lectal Variation in Constructional Semantics: Benefactive Ditransitives in Dutch. In: Dirk Geeraerts, Gitte Kristiansen & Yves Peirsman (eds.), *Advances in Cognitive Sociolinguistics*, 191-221. Berlin/New York: Mouton De Gruyter.
Colleman, Timothy. 2015. Constructionalization and Post-Constructionalization: The Constructional Semantics of the Dutch *Krijgen*-Passive from a Diachronic Perspective. In: Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic Construction Grammar*, 213-255. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Colleman, Timothy & Bernard De Clerck. 2011. Constructional Semantics on the Move: On Semantic Specialization in the English Double Object Construction. *Cognitive Linguistics* 22, 183-209.
Colleman, Timothy & Gunther De Vogelaer. 2003. De Benefactiefconstructie in De Zuidelijk-Nederlandse Dialecten. *Taal en Tongval* 15/16, 184-208.
Colleman, Timothy & Freek Van de Velde. 2015. Variatie En Verandering in Constructies. Op Het Snijvlak Van De Constructiegrammatica En De Variatielinguïstiek. *Taal en Tongval* 65 (2), 135-148.
Cotter, Colleen. 2003. Prescription and Practice. Motivations Behind Change in News Discourse. *Journal of Historical Pragmatics* 4, 45-74.
Coussé, Evie, Peter Andersson & Joel Olofsson. 2018. *Grammaticalization Meets Construction Grammar*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Croft, William. 1998. Linguistic Evidence and Mental Representation. *Cognitive Linguistics* 9 (2), 151-173.
Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William. 2003. Lexical Rules Vs. Constructions: A False Dichotomy. In: Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in Language: Studies in Honor of Günter Radden*, 49-68.
Croft, William & Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
Curzan, Anne. 2009. Historical Corpus Linguistics and Evidence of Language Change. In: Anke Lüdeling & Merja Kyto (eds.), *Corpus Linguistics: An International Handbook*, 1091-1109. Berlin/New York: Mouton De Gruyter.
D'Arcy, Alexandra. 2015. Stability, Stasis and Change. The Longue Durée of Intensification. *Diachronica* 32 (4), 449-493.
Dąbrowska, Ewa. 2004. *Language, Mind and Brain. Some Psychological and Neurological Constraints on Theories of Grammar*. Edinburgh: Edinburgh University Press.
Daems, Jocelyn, Kris Heylen & Dirk Geeraerts. 2013. 'Some Rise by Sin, and Some by Virtue Fall': Lexical Convergence between Belgian Dutch and Netherlandic Dutch. Paper presented at 5th Conference on Quantitative Investigations in Theoretical Linguistics, Leuven.
Dal, Georgette. 2003. Productivité Morphologique: Définitions Et Notions Connexes. In: Georgette Dal (ed.), *La Productivité Morphologique En Questions Et En Expérimentations*, vol. 140 (Langue Française), 3-23. Paris: Larousse.
De Clerck, Bernard & Timothy Colleman. 2013. From Noun to Intensifier: *Massa* and *Massa's* in Flemish Varieties of Dutch [Mar]. *Language Sciences* 36, 147-160.
De Clerck, Bernard et al. 2011. The Syntactic Flexibility of (New) Verbs of Instrument of Communication. *Functions of Language* 18 (1), 57-86.
De Cuypere, Ludovic. 2014. A Multivariate Analysis of the Old English Acc+Dat Double Object Alternation. *Corpus Linguistics and Linguistic Theory* 11 (2), 225-254.
De Smet, Hendrik. 2012. The Course of Actualization. *Language* 88, 601-633.
De Smet, Hendrik. 2013. *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford/New York: Oxford University Press.
De Smet, Hendrik. 2016. How Gradual Change Progresses: The Interaction between Convention and Innovation. *Language Variation and Change* 28, 83-102.
De Smet, Hendrik. 2017. *Entrenchment Effects in Language Change* (Entrenchment, Memory and Automaticity: The Psychology of Linguistic Knowledge and Language Learning). Berlin/New York: Mouton De Gruyter.
De Smet, Hendrik et al. 2018. The Changing Functions of Competing Forms. Attraction and Differentiation. *Cognitive Linguistics* 29 (2), 197-234.
De Smet, Hendrik & Olga Fischer. 2017. The Role of Analogy in Language Change. Supporting Constructions. In: Marianne Hundt (ed.), *The Changing English Language: Psycholinguistic Perspectives*, 240-268.
De Smet, Hendrik, Lobke Ghesquière & Freek Van de Velde. 2015. *On Multiple Source Constructions in Language Change*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
De Vooys, Cornelis Gerrit Nicolaas. 1975. *Geschiedenis Van De Nederlandse Taal*. Groningen: Tjeenk Willink.
Debrabandere, Peter. 2005. Verschillen Tussen Het Nederlandse En Het Belgische Nederlands. Paper presented at https://itidn.files.wordpress.com/2010/07/lezing-antwerpen-lessius-12-06-2005.pdf (last accessed on 1 September 2017).
Denison, David & Richard Hogg. 2006. *A History of the English Language*. Cambridge: Cambridge University Press.

Diessel, Holger. 2004. A Dynamic Network Model of Grammatical Constructions. In: Holger Diessel (ed.), *The Acquisition of Complex Sentences*, 13-40. Cambridge: Cambridge University Press.
Diessel, Holger. 2015. Usage-Based Construction Grammar. In: Ewa Dąbrowska & Dagmar Divjak (eds.), *Handbook of Cognitive Linguistics*, 295-321. Berlin/New York: Mouton De Gruyter.
Diessel, Holger & Martin Hilpert. 2016. Frequency Effects in Grammar. In: Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*. New York: Oxford University Press.
Diewald, Gabriele. 2002. A Model for Relevant Types of Contexts in Grammaticalization. In: Ilse Wischer & Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 103-120. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Doetjes, Jenny. 2008. Adjectives and Degree Modification. In: Louise McNally & Christopher Kennedy (eds.), *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, 123-155. Oxford: Oxford University Press.
Ellis, Nick C. & Fernando Ferreira-Junior. 2009. Constructions and Their Acquisitions: Islands and the Distinctiveness of Their Occupancy. *Annual Review of Cognitive Linguistics* 7, 187-220.
Erman, Britt & Beatrice Warren. 2000. The Idiom Principle and the Open Choice Principle. *Text* 20 (1), 29-62.
Espinal, Teresa M. & Jaume Mateu. 2010. Classes of Idioms and Their Interpretation. *Journal of Pragmatics* 42 (5), 1397-1411.
Everaert, Martin & Alexis Dimitriadis. 2013. On Reflexive Resultatives. Paper presented at Workshop Secondary Predication in Formal Frameworks, Utrecht University.
Felfe, Marc. 2012. Transitive Resultativkonstruktionen in Der Konstruktionsgrammatik. *Zeitschrift für germanistische Linguistik* 40 (3), 352-395.
Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone. *Language* 64 (3), 501-538.
Firth, John R. 1957. A Synopsis of Linguistic Theory 1930-1955. In: John R. Firth (ed.), *Studies in Linguistic Analysis*. Oxford: Blackwell.
Fried, Mirjam. 2013. Principles of Constructional Change. In: Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 419-437. Oxford: Oxford University Press.
Fries, Charles C. 1940. *American English Grammar. The Grammatical Structure of Present-Day American English with Especial Reference to Social Differences or Class Dialects*. New York: Appleton-Century-Crofts, Inc.
Gaeta, Livio & Davide Ricca. 2006. Productivity in Italian Word Formation: A Variable-Corpus Approach. *Linguistics* 44 (1), 57-89.
Geeraerts, Dirk. 1993. Vagueness's Puzzles, Polysemy's Vagaries. *Cognitive Linguistics* 4 (3), 223-272.
Geeraerts, Dirk, Stefan Grondelaers & Dirk Speelman. 1999. *Convergentie En Divergentie in De Nederlandse Woordenschat. Een Onderzoek Naar Kleding- En Voetbaltermen*. Amsterdam: PJ Meertens Instituut.
Geeraerts, Dirk, Gitte Kristiansen & Yves Peirsman. 2010. *Advances in Cognitive Sociolinguistics*. Berlin/New York: Mouton De Gruyter.
Geleyn, Tim. 2016. *Constructies in Variatie En Verandering. Diachroon Corpusonderzoek Naar De Nederlandse Aan-Constructie Vanuit Semasiologisch En Onomasiologisch Perspectief*. PhD dissertation, Universiteit Gent.
Geurts, Bart. 2000a. Explaining Grammaticalization (the Standard Way). *Linguistics* 38 (4), 781-788.

Geurts, Bart. 2000b. Function or Fashion? Reply to Martin Haspelmath. *Linguistics* 38 (4), 1175-1180.
Gibbs, Raymond W. et al. 1989. Speakers' Assumptions About the Lexical Flexibility of Idioms. *Memory and Cognition* 16, 58-68.
Glasbey, Sheila. 2003. Let's Paint the Town Red for a Few Hours: Composition of Aspect in Idioms. In: A.M. Wellington (ed.), *Proceedings of the Acl Workshop: The Lexicon and Figurative Language*, 42-48. Sapporo:
Glynn, Dylan. 2010. Testing the Hypothesis: Objectivity and Verification in Usage-Based Cognitive Semantics. In: Dylan Glynn & Kerstin Fischer (eds.), *Quantitative Methods in Cognitive Semantics. Corpus-Driven Approaches*, 239-270. Berlin/New York: Mouton De Gruyter.
Goldberg, Adele E. 1993. Another Look at Some Learnability Paradoxes. In: Eve V. Clark (ed.), *Proceedings of the Twenty-Fifth Annual Child Language Research Forum*, 60-75. Stanford, CA: CSLI Publications.
Goldberg, Adele E. 1995. *Constructions : A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Goldberg, Adele E. 2003. Constructions: A New Theoretical Approach to Language. *Trends in Cognitive Science* 7 (5), 219-224.
Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford/New York: Oxford University Press.
Goldberg, Adele E. 2011. Corpus Evidence of the Viability of Statistical Preemption. *Cognitive Linguistics* 22 (1), 131-153.
Goldberg, Adele E. 2013. Constructionist Approaches. In: Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 15-31. Oxford: Oxford University Press.
Goldberg, Adele E., Devin M. Casenhiser & Nitya Sethuraman. 2004. Learning Argument Structure Generalizations. *Cognitive Linguistics* 15 (3), 289-316.
Goldberg, Adele E. & Ray Jackendoff. 2004. The English Resultative as a Family of Constructions. *Language* 80 (3), 532-568.
Greenbaum, Sidney. 1970. *Verb-Intensifier Collocations in English. An Experimental Approach*. Den Haag: Mouton.
Gries, Stefan Th. 2007. *Coll.Analysis 3.2a. A Program for R for Windows 2.X*.
Gries, Stefan Th. 2012. Frequencies, Probabilities, and Association Measures in Usage-/Exemplar-Based Linguistics. *Studies in Language* 11 (3), 477-510.
Gries, Stefan Th. 2013a. 50-Something Years of Work on Collocations. What Is or Should Be Next... *International Journal of Corpus Linguistics* 18 (1), 137-165.
Gries, Stefan Th. 2013b. *Statistics for Linguistics with R: A Practical Introduction (2nd Edition)*. Berlin/New York: Mouton De Gruyter.
Gries, Stefan Th. 2015. More (Old and New) Misunderstandings of Collostructional Analysis: On Schmid and Küchenhoff (2013). *Cognitive Linguistics* 26 (3), 505-536.
Gries, Stefan Th. & Martin Hilpert. 2008. The Identification of Stages in Diachronic Data: Variability-Based Neighbor Clustering. *Corpora* 3 (1), 59-81.
Gries, Stefan Th. & Martin Hilpert. 2012. Variability-Based Neighbor Clustering: A Bottom-up Approach to Periodization in Historical Linguistics. In: T. Nevalainen & E. C. Traugott (eds.), *The Oxford Handbook of the History of English*, 134-144. Oxford: Oxford University Press.
Gries, Stefan Th. & Anatol Stefanowitsch. 2004a. Covarying Collexemes in the *into*-Causative. In: Michel Achard & Suzanne Kemmer (eds.), *Language, Culture, and Mind*, 225-236. Stanford, CA: CSLI Publications.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004b. Extending Collostructional Analysis: A Corpus-Based Perspective on 'Alternations'. *International Journal of Corpus Linguistics* 9 (1), 97-129.
Gries, Stefan Th. & Anatol Stefanowitsch. 2010. Cluster Analysis and the Identification of Collexeme Classes. In: Sally Rice & John Newman (eds.), *Empirical and Experimental Methods in Cognitive/Functional Research*, 73-90. Stanford, CA: CSLI Publications.
Gries, StefanTh. 2006. Corpus-Based Methods and Cognitive Semantics: The Many Meanings of *to Run*. In: Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*, 57-99. Berlin/New York: Mouton De Gruyter.
Grondelaers, Stefan, Dirk Speelman & Dirk Geeraerts. 2008. National Varieties in the Use of *Er* "There": Regional and Diachronic Constraints on Cognitive Explanations. In: Gitte Kristiansen & René Dirven (eds.), *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*, 153-203. Berlin/New York: Mouton De Gruyter.
Grondelaers, Stefan et al. 2001. Inhoudswoorden En Preposities Als Standaardiserings-Indicatoren. De Diachrone En Synchrone Status Van Het Belgische Nederlands. *Nederlandse Taalkunde* 6, 179-202.
Gutzmann, Daniel & Katharina Turgay. 2012. Expressive Intensifiers in German: Syntax-Semantics Mismatches. In: Christopher Piñon (ed.), *Empirical Issues in Syntax and Semantics*, vol. 9, 149-166. Exclusively online: http://www.cssp.cnrs.fr/eiss9/.
Haeseryn, Walter. 1996. Grammaticale Verschillen Tussen Het Nederlands in België En Het Nederlands in Nederland: Een Poging to Inventarisatie. In: Roeland Van Hout, Joep Kruijsen & Toon Hagen (eds.), *Taalvariaties. Toonzettingen En Modulaties Op Een Thema*. Dordrecht: Foris.
Hanks, Patrick. 2004. The Syntagmatics of Metaphor and Idiom. *International Journal of Lexicography* 17 (3), 245-274.
Harder, Peter. 2010. *Meaning in Mind and Society: A Functional Contribution to the Social Turn in Cognitive Linguistics*. Berlin/New York: Mouton De Gruyter.
Haspelmath, Martin. 1999. Why Is Grammaticalization Irreversible. *Linguistics* 37 (6), 1043-1068.
Haspelmath, Martin. 2000. The Relevance of Extravagance: A Reply to Bart Geurts. *Linguistics* 38 (4), 789-798.
Heine, Bernd. 2002. On the Role of Context in Grammaticalization. In: Ilse Wischer & Gabriele Diewald (eds.), *New Reflections on Grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Heylen, Kris et al. 2015. Monitoring Polysemy: Word Space Models as a Tool for Large-Scale Lexical Semantic Analysis. *Lingua* 157, 153-172.
Hilpert, Martin. 2011. Diachronic Collostructional Analysis: How to Use It and How to Deal with Confounding Factors. In: Kathryn Allan & Justyna A. Robinson (eds.), *Current Methods in Historical Semantics*, 130-160. Berlin/New York: Mouton De Gruyter.
Hilpert, Martin. 2012. Diachronic Collostructional Analysis Meets the Noun Phrase: Studying *Many a* Noun in Coha. In: T. Nevalainen & E. C. Traugott (eds.), *The Oxford Handbook of the History of English*, 233-244.
Hilpert, Martin. 2013. *Constructional Change in English : Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.
Hilpert, Martin. 2014. *Construction Grammar and Its Application to English*. Edinburgh: Edinburgh University Press.
Hilpert, Martin. 2015a. From *Hand-Carved* to *Computer-Based*: Noun-Participle Compounding and the Upward Strengthening Hypothesis. *Cognitive Linguistics* 26 (1), 113-147.

Hilpert, Martin. 2015b. Historical Linguistics. In: Ewa Dąbrowska & Dagmar Divjak (eds.), *Handbook of Cognitive Linguistics*. Berlin/New York: Mouton De Gruyter.
Hilpert, Martin. 2017. Frequencies in Diachronic Corpora and Knowledge of Language. In: Marianne Hundt, Simone Pfenninger & Sandra Mollin (eds.), *The Changing English Language - Psycholinguistic Perspectives*, 49-68. Cambridge: Cambridge University Press.
Hilpert, Martin. 2018. Three Open Questions in Diachronic Construction Grammar. In: Evie Coussé, Peter Andersson & Joel Olofsson (eds.), *Grammaticalization Meets Construction Grammar*, vol. 21-39. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Hilpert, Martin & Holger Diessel. 2017. Entrenchment in Construction Grammar. *Entrenchment and the Psychology of Language Learning. How We Reorganize and Adapt Linguistic Knowledge*, 57-74. Berlin/New York: Mouton De Gruyter.
Hilpert, Martin & Stefan Th. Gries. 2009. Assessing Frequency Changes in Multistage Diachronic Corpora: Applications for Historical Corpus Linguistics and the Study of Language Acquisition. *Literary and Linguistic Computing* 24 (4), 385-401.
Himmelmann, Nikolaus. 2004. Lexicalization and Grammaticization: Opposite or Orthogonal? In: Walter Bisang, Nikolaus Himmelmann & Björn Wiener (eds.), *What Makes Grammaticalization? A Look from Its Fringes and Its Components* (Trends in Linguistics, Studies and Monographs), 19-40. Berlin/New York: Mouton De Gruyter.
Hockett, Charles F. 1958. *A Course in Modern Linguistics*. New York: The Macmillan.
Hoeksema, Jack. 2005. Rijkdom En Weelde Van Het Nederlands. *TABU* 34 (1/2), 1-12.
Hoeksema, Jack. 2012. Elative Compounds in Dutch: Properties and Developments. In: Guido Oebel (ed.), *Intensivierungskonzepte Bei Adjektiven Und Adverbien Im Sprachvergleich [Cross-Linguistic Comparison of Intensified Adjectives and Adverbs]*, 97-142. Hamburg: Verlag dr. Kovač.
Hoeksema, Jack & Jan Korterink. 2011. Keus Van Bijwoorden Van Hoge Graad: Een Vergelijking Van Studenten Met Vmbo-Leerlingen. *TABU* 39 (1/2), 23-37.
Hoeksema, Jack & Donna Jo Napoli. 2008. Just for the Hell of It: A Comparison of Two Taboo-Term Constructions. *Journal of Linguistics* 44 (2), 347-378.
Hoekstra, Teun. 1988. Small Clause Result. *Lingua* 74, 101-139.
Hoekstra, Teun. 1992. Small Clause Theory. *Belgian Journal of Linguistics* 7, 125-151.
Hoffmann, Sebastian & Joybrato Mukherjee. 2007. Ditransitive Verbs in Indian English and British English: A Corpus-Linguistic Study. *Arbeiten aus Anglistik und Amerikanistik* 32, 5-24.
Hoffmann, Thomas & Graeme Trousdale. 2013. *The Oxford Handbook of Construction Grammar*. Oxford/New York: Oxford University Press.
Hoppenbrouwers, Cor. 1991. *Jongerentaal. De Tipparade Van De Omgangstaal.* Hoogezand: Stubeg.
Hopper, Paul J. 1991. On Some Principles of Grammaticization. In: E. C. Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization: Volume I. Theoretical and Methodological Issues*, 17-36. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
Howe, Nicholas. 1988. Metaphor in Contemporary American Political Discourse. *Metaphor and Symbolic Activity* 3 (2), 87-104.
Huang, James. 2006. Resultatives and Unaccusatives: A Parametric View. *Bulletin of the Chinese Linguistic Society of Japan* 2006 (253), 1-43.
Hübler, Axel. 1998. *The Expressivity of Grammar. Grammatical Devices Expressing Emotion across Time* (Topics in English Linguistics). Berlin/New York: Mouton De Gruyter.
Hudson, Richard. 2007. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press.

Israel, Michael. 1996. The *Way* Constructions Grow. In: Adele E. Goldberg (ed.), *Conceptual Structure, Discourse and Language*, 217-230. Stanford, CA: CSLI Publications.
Ito, Rika & Sali Tagliamonte. 2003. *Well Weird*, *Right Dodgy*, *Very Strange*, *Really Cool*: Layering and Recycling in English Intensifiers. *Language in Society* 32 (2), 257-279.
Iwata, Seizi. 2006. Argument Resultatives and Adjunct Resultatives in a Lexical Constructional Account: The Case of Resultatives with Adjectival Result Phrases. *Language Sciences* 28 (5), 449-496.
Iwata, Seizi. 2008. *Locative Alternation: A Lexical-Constructional Approach*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Jackendoff, Ray. 1997. Twisting the Night Away. *Language* 73, 534-559.
Jackendoff, Ray. 2008. Construction after Construction and Its Theoretical Challenges. *Language* 84 (1), 8-28.
Jackendoff, Ray. 2013. Constructions in the Parallel Architecture. In: Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 70-92. Oxford: Oxford University Press.
Janda, Richard D. 2001. Beyond "Pathways" and "Unidirectionality": On the Discontinuity of Language Transmission and the Counterability of Grammaticalization. *Language Sciences* 23 (2-3), 265-340.
Jespersen, Otto H. 1922. *Language: Its Nature, Development, and Origin*. London: George Allen & Unwin.
Jing-Schmidt, Zhuo. 2007. Negativity Bias in Language: A Cognitive-Affective Model of Emotive Intensifiers. *Cognitive Linguistics* 18 (3), 417-443.
Joseph, Brian D. 2001. Is There Such a Thing as 'Grammaticalization'? *Language Sciences* 23, 163-186.
Jurafsky, Daniel. 1992. *An on-Line Computational Model of Human Sentence Interpretation: A Theory of the Representation and Use of Linguistic Knowledge*. Ph.D. dissertation, University of California.
Kay, Paul. 2013. The Limits of (Construction) Grammar. In: Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Contruction Grammar*, 32-48. Oxford: Oxford University Press.
Kay, Paul & Charles John Fillmore. 1999. Grammatical Constructions and Linguistic Generalizations: The What's X Doing Y? Construction. *Language* 72 (1), 1-33.
King, Ruth. 2013. Expressivity and Innovation. In: Ruth King (ed.), *Acadian French in Time and Space: A Study in Morphosyntax and Comparative Sociolinguistics*, vol. 97 (Publication of the American Dialect Society), 99-111. Durham, NC: Duke University Press.
Kipper-Schuler, Karin. 2006. *Verbnet: A Broad-Coverage, Comprehensive Verb Lexicon.* PhD dissertation, University of Pennsylvania.
Kiss, Katalin É. 2006. The Function and the Syntax of the Verbal Particle. In: Katalin É. Kiss (ed.), *Event Structure and the Left Periphery: Studies on Hungarian*, 17-55. Dordrecht: Springer.
Kitis, Eliza & Michalis Milapides. 1997. Read It and Believe It: How Metaphor Constructs Ideology in News Discourse. A Case Study. *Journal of Pragmatics* 28, 557-590.
Klein, Henny. 1998. *Adverbs of Degree in Dutch and Related Languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Kristiansen, Gitte & René Dirven. 2008. *Cognitive Sociolinguistics. Language Variation, Cultural Models, Social Systems*. Berlin/New York: Mouton De Gruyter.
Kudo, Shun. 2011. A Comparative Study between Resultative Constructions and Body-Part-*Off* Constructions. *Tsubaka English Studies* 29, 169-185.
Kuryłowicz, Jerzy. 1975. The Evolution of Grammatical Categories. In: Jerzy Kuryłowicz (ed.), *Esquisses Linguistiques*, 38-54. Munich: Wilhelm Fink Verlag.

Lakoff, George & Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
Landauer, Thomas K., Peter W. Foltz & Darrell Laham. 1998. An Introduction to Latent Semantic Analysis. *Discourse Processes* 25 (259-284),
Landauer, Thomas K. et al. 2007. *Handbook of Latent Semantic Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
Landsbergen, Frank. 2009. *Cultural Evolutionary Modeling of Patterns in Language Change: Exercises in Evolutionary Linguistics*. PhD dissertation, Leiden University.
Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*, vol. 1. Stanford, CA: Stanford University Press.
Langacker, Ronald W. 1999. *Grammar and Conceptualization*. Berlin/New York: Mouton De Gruyter.
Langacker, Ronald W. 2005. Construction Grammars: Cognitive, Radical and Less So. In: Francisco J. Ruiz de Mendoza Ibáñez & Sandra Peña Cervel (eds.), *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*, 101-159. Berlin/New York: Mouton De Gruyter.
Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
Lauwers, Peter & Dominique Willems. 2011. Coercion: Definition Adn Challenges, Current Approaches and New Trends. *Linguistics* 49 (6), 1219-1235.
Lehmann, Christian. 1995 [1982]. *Thoughts on Grammaticalization*. München/Newcastle: Lincom Europe.
Lehmann, Christian. 2002. New Reflections on Grammaticalization and Lexicalization. In: Ilse Wischer & Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 1-18. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago: The University of Chicago Press.
Levshina, Natalia, Dirk Geeraerts & Dirk Speelman. 2013. Towards a 3d-Grammar: Variation in the Use of Dutch Causative Constructions. *Journal of Pragmatics* 52, 34-48.
Levshina, Natalia, Dirk Geeraerts & Dirk Speelman. 2014. Dutch Causative Constructions with *Doen* and *Laten*: Quantification of Meaning and Meaning of Quantification. In: D. Glynn & J. Robinson (eds.), *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*, 377-414. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Lieber, Rochelle. 1992. *Deconstruction Morphology: Word Formation in Syntactic Theory*. Chicago, IL: University of Chicago Press.
Lieven, Elena & Michael Tomasello. 2008. Children's First Language Acquistion from a Usage-Based Perspective. In: Peter Robinson & Nick C. Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Taylor and Francis.
Lorenz, David. 2012. *Contractions of English Semi-Modals: The Emancipating Effect of Frequency*. Unpublished PhD dissertation, University of Freiburg.
Lorenz, Gunter. 2002. *Really Worthwhile* or N*ot Really Significant*? A Corpus-Based Approach to Delexicalization and Grammaticalization of Intensifiers in Modern English. In: Ilse Wischer & Gabriele Diewald (eds.), *New Reflections on Grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Luzondo-Oyon, Alba. 2014. Constraining Factors on the Family of Resultative Constructions. *Review of Cognitive Linguistics* 12 (1), 30-63.
Mair, Christian. 2006. *Twentieth-Century English: History, Variation and Standardization*. Cambridge: Cambridge University Press.

Margerie, Hélène. 2011. Grammaticalising Constructions: *To Death* as a Peripheral Degree Modifier. *Folia Linguistica Historica* 32 (1), 115-147.
Margerie, Hélène. 2013. Can You Literally Be Scared Sick? The Role of Analogy in the Rise of a Network of Resultative and Degree Modifier Constructions. In: Anna Giacalone Ramat, Caterina Mauri & Piera Molinelli (eds.), *Synchrony and Diachrony: A Dynamic Interface*, 79-104. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Margerie, Hélène. 2014. He Was Angry Awful: Intertwining Paths of Development to New Degree Modifier Constructions in American English. *American Speech* 89 (3), 257-287.
Markham, Tim. 2012. The Politics of Journalistic Creativity: Expressiveness, Authenticity and De-Authorization. *Journalism Practice* 6 (2), 187-200.
McCarthy, Michael & Ronald Carter. 2004. There's Millions of Them: Hyperbole in Everyday Conversation. *Journal of Pragmatics* 36 (2), 149-184.
McGlone, Matthew S., Sam Glucksberg & Cristina Cacciari. 1994. Semantic Productivity and Idiom Comprehension [Mar-Apr]. *Discourse Processes* 17 (2), 167-190.
Meillet, Antoine. 1912. *L'évolution Des Formes Grammaticales*. Bologna/Paris/London: N.Zanichelli/F.Alcan /Williams et Norgate.
Méndez-Naya, Belén. 2003. On Intensifiers and Grammaticalization: The Case of *Swithe*. *English Studies* 84 (4), 372-391.
Michaelis, Laura A. 2002. Headless Constructions and Coercion by Construction. In: E. Francis & Laura A. Michaelis (eds.), *Mismatch: Form-Function Incongruity and the Architecture of Grammar*, 259-310. Stanford, CA: CSLI Publications.
Michaelis, Laura A. 2004. Implicit and Explicit Type-Shifting in Construction Grammar. *Cognitive Linguistics* 15 (1), 1-67.
Mukherjee, Joybrato & Sebastian Hoffmann. 2006. Describing Verb-Complementational Profiles of New Englishes: A Pilot Study of Indian English. *English World-Wide* 27, 147-173.
Musolff, Andreas. 2010. Political Imagery of Europa: A House without Exit Doors? *Journal of Multilingual and Multicultural Development* 21 (3), 216-229.
Napoli, Donna Jo & Jack Hoeksema. 2009. The Grammatical Versatility of Taboo Terms. *Studies in Language* 33 (3), 612-643.
Narrog, Heiko & Bernd Heine. 2011. *The Oxford Handbook of Grammaticalization*. Oxford/New York: Oxford University Press.
Neeleman, Ad, Hans Van de Koot & Jenny Doetjes. 2004. Degree Expressions. *The Linguistic Review* 21, 1-66.
Neveu, Erik. 2014. Revisiting Narrative Journalism as One of the Futures of Journalism. *Journalism Studies* 15 (5), 533-542.
Noël, Dirk. 2007. Diachronic Construction Grammar and Grammaticalization Theory. *Functions of Language* 14 (2), 177-202.
Norde, Muriel. 2006. Van Suffix Tot Telwoord Tot Bijwoord: Degrammaticalisering En (Re)Grammaticalisering Van *Tig*. *TABU* 35 (1-2), 33-60.
Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.
Norde, Muriel, Bernard De Clerck & Timothy Colleman. 2014. The Emergence of Non-Canonical Degree Modifiers in Non-Standard Varieties of Dutch: A Constructionalization Perspective. In: Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*, 207-250. Berlin/New York: Mouton De Gruyter.
Norde, Muriel & Oscar Strik. 2017. Analogy, Multiple Sources and Constructional Networks: The Dutch Klukkluk Construction. In: Anne Bannink & Wim Honselaar (eds.), *From Variation to Iconicity: Festschrift for Olga Fischer*, 295-310. Amsterdam: Pegasus.

Norde, Muriel & Kristel Van Goethem. 2014. Bleaching, Productivity and Debonding of Prefixoids. A Corpus-Based Analysis of 'Giant' in German and Swedish. *Linguisticae Investigationes* 37 (2), 256-274.
Offstein, Evan H. & Christopher P. Neck. 2003. From "Acing the Test" to "Touching Base": The Sports Metaphor in the Classroom. *Business and Professional Communication Quarterly* 66 (4), 23-35.
Ogarkova, A. 2007. "Green-Eyed Monsters": A Corpus-Based Study of Metaphoric Conceptualizations of Jealousy and Envy in Modern English. *Metaphorik* 13, 87-147.
Ogura, Mikeo & William S-Y. Wang. 1996. Snowball Effects in Lexical Diffusion. The Development of *-S* in the Third Person Singular Present Indicative in English. In: Derek Britton (ed.), *English Historical Linguistics 1994. Papers from the 8th International Conference on English Historical Linguistics*, 119-141. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Onysko, Alexander & Andreea S. Calude. 2014. Comparing the Usage of Māori Loans in Spoken and Written New Zealand English: A Case Study of Māori, Pākeha, and Kiwi. In: Eline Zenner & Gitte Kristiansen (eds.), *New Perspectives on Lexical Borrowing: Onomasiological, Methodological and Phraseological Innovations*, 143-169. Berlin/New York: Mouton De Gruyter.
Oostdijk, Nelleke et al. 2013. The Construction of a 500-Million-Word Reference Corpus of Contemporary Written Dutch. In: Peter Spyns & Jan Odijk (eds.), *Essential Speech and Language Technology for Dutch: Results by the Stevin-Programme*. Berlin: Springer.
Paradis, Carita. 2008. Configurations, Construals and Change: Expressions of Degree. *English Language and Linguistics* 12 (2), 317-343.
Partington, Alan. 1993. Corpus Evidence of Language Change: The Case of the Intensifier. In: Mona Baker, Gill Francis & Elena Tognini-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*, 177-192. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Pasma, Trijntje. 2011. *Metaphor and Register Variation. The Personalization of Dutch News Discourse.* PhD dissertation, Vrije Universiteit Amsterdam.
Pearce, Michael. 2005. Informalization in Uk Party Election Broadcasts: 1966-97. *Language & Literature* 14 (1), 65-90.
Peña-Cervel, Maria S. 2016. Argument Structure and Implicational Constructions at the Crossroads. *Review of Cognitive Linguistics* 14 (2), 474-497.
Perek, Florent. 2015. *Argument Structure in Usage-Based Construction Grammar. Experimental and Corpus-Based Perspectives*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Perek, Florent. 2016a. Recent Change in the Productivity and Schematicity of the Way-Construction: A Distributional Semantic Analysis. *Corpus Linguistics and Linguistic Theory* Published ahead of print,
Perek, Florent. 2016b. Using Distributional Semantics to Study Syntactic Productivity in Diachrony: A Case Study. *Linguistics* 54 (1), 149-188.
Perek, Florent & Adele E. Goldberg. 2017. Linguistic Generalization on the Basis of Function and Constraints on the Basis of Statistical Preemption. *Cognition* 168, 276-293.
Perek, Florent & Martin Hilpert. 2017. A Distributional Semantic Approach to the Periodization of Change in the Productivity of Constructions. *International Journal of Corpus Linguistics* 22 (4), 490-520.
Pertejo, Paloma N. & Ignacio M. Palacios Martínez. 2014. That's Absolutely Crap, Totally Rubbish. The Use of the Intensifiers Absolutely and Totally in the Spoken Language of British Adults and Teenagers. *Functions of Language* 21 (2), 210-237.

Peters, Hans. 1994. Degree Adverbs in Early Modern English. In: Dieter Kastovsky (ed.), *Studies in Early Modern English*, 269-288. Berlin/New York: Mouton De Gruyter.
Pijpops, Dirk & Freek Van de Velde. 2016. Constructional Contamination: How Does It Work and How Do We Measure It? *Folia Linguistica* 50 (2), 543-581.
Pit, Mirna. 2003. *How to Express Yourself with a Causal Connective*. PhD dissertation, Universiteit Utrecht.
Plag, Ingo. 1999. *Morphological Productivity. Structural Constraints in English Derivation*. Berlin/New York: Mouton De Gruyter.
Plag, Ingo, Christiane Dalton-Puffer & Harald Baayen. 1999. Morphological Productivity across Speech and Writing. *English Language and Linguistics* 3 (2), 209-228.
Pütz, Martin, Justyna A. Robinson & Monika Reif. 2012. The Emergence of Cognitive Sociolinguistics: An Introduction. *Review of Cognitive Linguistics* 10, 241-263.
Quirk, Randolph et al. 1985. *A Comprehensive Grammar of the English Language*. New York: Longman.
Rainer, Franz. 2015. Intensification. In: Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-Formation: An International Handbook of the Languages of Europe*, vol. 2, Ch 77. Berlin/New York: Mouton De Gruyter.
Rappaport Hovav, Malka & Beth Levin. 1998. Building Verb Meaning. In: M. Butt & W. Geuder (eds.), *The Projection of Arguments*, 97-134. Stanford, CA: CSLI Publications.
Rappaport Hovav, Malka & Beth Levin. 2001. An Event Structure Account of English Resultatives. *Language* 77, 766-797.
Redington, Martin, Nick Chater & Steven Finch. 1998. Distributional Information: A Powerful Cue for Acquiring Syntactic Categories. *Cognitive Science* 22 (4), 435-469.
Reker, Siemon. 1996. *Dikke Woorden. Bikkelhaard, Bragelvet, Strontdeurnat En Hun Soortgenoten in Het Gronings En Verwante Talen Verzameld, Geordend, Toegelicht En Verklaard.* Bedum: Profiel Uitgeverij.
Robertson, Stuart & Frederic G. Cassidy. 1954. *The Development of Modern English*. Englewood Cliffs, NJ: Prentice-Hall.
Roggenkamp, Karen. 2005. *Narrating the News: New Journalism and Literary Genre in Late Nineteenth-Century American Newspapers and Fiction*. Kent: Kent State University Press.
Rosemeyer, Malte. 2014. *Auxiliary Selection in Spanish. Gradience, Gradualness, and Conservation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Santa Ana, Otto. 1999. `Like an Animal I Was Treated': Anti-Immigrant Metaphor in Us Public Discourse. *Discourse & Society* 10 (2), 191-224.
Sawada, Shigeyasu. 2000. The Semantics of the 'Body Part *Off*' Construction. *English Linguistics* 17 (2), 361-385.
Scheibman, Joanne. 2002. *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Schmid, Hans-Jörg. 2000. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin/New York: Mouton De Gruyter.
Schmid, Hans-Jörg. 2010. Does Frequency in Text Instantiate Entrenchment in the Cognitive System? In: Dylan Glynn & Kerstin Fischer (eds.), *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, 101-134. Berlin/New York: Mouton De Gruyter.
Schmid, Hans-Jörg. 2017. A Framework for Understanding Linguistic Entrenchment and Its Psychological Foundations. In: Hans-Jörg Schmid (ed.), *Entrenchment and the Psychology of Language Learning. How We Reorganize and Adapt Linguistic Knowledge*, 9-36. Berlin/New York: Mouton De Gruyter.

Schneider-Wiejowski, Karina. 2009. Sprachwandel Anhand Von Produktivitätsverschiebungen in Der Schweizerdeutschen Derivationsmorphologie *Linguistik Online* 38 (2), 79-90.
Schröder, P.H. 1980. *Van Aalmoes Tot Zwijntjesjager*. Vijverhof, Baarn: Erven Thomas Rap.
Schultink, Hendrik. 1961. Produktiviteit Als Morfologisch Fenomeen. *Forum der Letteren* 2, 100-125.
Schultink, Hendrik. 1962. *De Morfologische Valentie Van Het Ongelede Adjectief in Modern Nederlands*. Den Haag: Van Goor Zonen.
Semino, Elena & Michela Masci. 1996. Politics Is Football: Metaphor in the Discourse of Silvio Berlusconi in Italy. *Discourse & Society* 7 (2), 243-269.
Sharpe, Donald. 2015. Your Chi-Square Test Is Statistically Significant: Now What? *Practical Assessment, Research & Evaluation* 20 (8), Available online: http://pareonline.net/getvn.asp?v=20&n=28.
Siewierska, Anna & Willem Hollmann. 2007. Ditransitive Clauses in English with Special Reference to Lancashire Dialect. In: Make Hannay & Gerard Steen (eds.), *Structural-Functional Studies in English Grammar*, 83-102. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Simpson, Jane. 1983. Resultatives. In: Lori Levin, Malka Rappaport & Annie Zaenen (eds.), *Papers in Lexical-Functional Grammar*, 143-157. Bloomington: Indiana University Linguistics Club.
Soriano, Cristina & Javier Valenzuela. 2009. Emotion and Colour across Languages: Implicit Associations in Spanish Colour Terms. *Social Science Information* 48 (3), 421-445.
Speelman, Dirk & Dirk Geeraerts. 2009. Causes for Causatives: The Case of Dutch *Doen* and *Laten*. In: Ted Sanders & Eve Sweetser (eds.), *Causal Categories in Discourse and Cognition*, 173-204. Berlin/New York: Mouton De Gruyter.
Speelman, Dirk, Stefan Grondelaers & Dirk Geeraerts. 2008. Variation in the Choice of Adjectives in the Varieties of Dutch. In: Gitte Kristiansen & René Dirven (eds.), *Cognitive Sociolinguistics. Language Variation, Cultural Models, Social Systems.*, 205-233. Berlin/New York: Mouton De Gruyter.
Steen, Gerard J. 2003. Conversationalization in Discourse: Stylistic Changes in Editorials of *the Times* between 1950 and 2000. In: Luuk Lagerwerf, Wilbert Spooren & Liesbeth Degand (eds.), *Determination of Information and Tenor in Texts: Multidisciplinary Approaches to Discourse*. Münster: Nodus Publikationen.
Steen, Gerard J. et al. 2010. *A Method for Linguistic Metaphor Identification from Mip to Mipvu*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Stefanowitsch, Anatol. 2008. Negative Entrenchment: A Usage-Based Approach to Negative Evidence. *Cognitive Linguistics* 19 (3), 513-531.
Stefanowitsch, Anatol & Susanne Flach. 2017. The Corpus-Based Perspective on Entrenchment. In: Hans-Jörg Schmid (ed.), *Entrenchment and the Psychology of Language Learning. How We Reorganize and Adapt Linguistic Knowledge*, 101-127. Berlin/New York: Mouton De Gruyter.
Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the Interaction of Words and Constructions. *International Journal of Corpus Linguistics* 8 (2), 209-243.
Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying Collexemes. *Corpus Linguistics and Linguistic Theory* 1 (1), 1-43.
Stenström, Anna-Brita. 1999. *He Was Really Gormless - She's Bloody Crap*: Girls, Boys and Intensifiers. In: Hilde Hasselgård & Signe Okesfjell (eds.), *Out of Corpora: Studies in Honour of Stig Johansson*, 69-77. Amsterdam: Rodopi.

Stoffel, Cornelis. 1901. *Intensives and Down-Toners: A Study in English Adverbs* (Anglistische Forschungen). Heidelberg: C. Winter's Universitätsbuchhandlung.
Suttle, Laura & Adele E. Goldberg. 2011. The Partial Productivity of Constructions as Induction. *Linguistics* 49 (6), 1237-1269.
Szmrecsanyi, Benedikt. 2010. The English Genitive Alternation in a Cognitive Sociolinguistics Perspective. In: Dirk Geeraerts, Gitte Kristiansen & Yves Peirsman (eds.), *Advances in Cognitive Sociolinguistics*, 143-166. Berlin/New York: Mouton De Gruyter.
Szmrecsanyi, Benedikt. 2016. About Text Frequencies in Historical Linguistics: Disentangling Environmental and Grammatical Change. *Corpus Linguistics and Linguistic Theory* 12 (1), 153-171.
Tagliamonte, Sali. 2008. So Different and Pretty Cool! Recycling Intensifiers in Toronto, Canada. *English Language and Linguistics* 12 (2), 361-394.
Taylor, John R. 2004. The Ecology of Constructions. In: Günter Radden & Klaus-Uwe Panther (eds.), *Studies in Linguistic Motivation*, 49-73. Berlin/New York: Mouton De Gruyter.
Ten Buuren, Maaike et al. to appear. Facking Nice! Een Onderzoek Naar De Intensiteit Van Intensiveerders. *Nederlandse Taalkunde*,
Tomasello, Michael. 1998. The Return of Constructions. *Journal of Child Language* 25 (2), 431-442.
Torrent, Tiago T. 2015. On the Relation between Inheritance and Change: The Constructional Convergence and the Construction Network Reconfiguration Hypotheses. In: Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic Construction Grammar*, 173-211. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Traugott, Elizabeth C. 1989. On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change. *Language* 57, 33-65.
Traugott, Elizabeth C. 2003. From Subjectification to Intersubjectification. In: R. Hickey (ed.), *Motives for Language Change*, 124-139. Cambridge: Cambridge University Press.
Traugott, Elizabeth C. 2007. The Concepts of Constructional Mismatch and Type-Shifting from the Perspective of Grammaticalization. *Cognitive Linguistics* 18 (4), 523-557.
Traugott, Elizabeth C. 2008a. The Grammaticalization of Np of Np Constructions. In: Alexander Bergs & Gabriele Diewald (eds.), *Constructions and Language Change*, 21-43. Berlin/New York: Mouton De Gruyter.
Traugott, Elizabeth C. 2008b. Grammaticalization, Constructions and the Incremental Development of Language: Suggestions from the Development of Degree Modifiers in English. In: Regine Eckardt, Gerhard Jäger & Tonjes Veenstra (eds.), *Variation, Selection, Development: Probing the Evolutionary Model of Language Change* (Trends in Linguistics, Studies and Monographs), 219-250. Berlin/New York: Mouton De Gruyter.
Traugott, Elizabeth C. 2010. (Inter)Subjectivity and (Inter)Subjectification: A Reassessment. In: Kristin Davidse, L. Vandelanotte & H. Kuyckens (eds.), *Subjectification, Intersubjectification and Grammaticalization*, 29-71. Berlin/New York: Mouton De Gruyter.
Traugott, Elizabeth C. & Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
Traugott, Elizabeth C. & Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
Trousdale, Graeme. 2008a. A Constructional Approach to Lexicalization Processes in the History of English: Evidence from Possessive Constructions. *Word Structure* 1 (2), 156-177.

Trousdale, Graeme. 2008b. Constructions in Grammaticalization and Lexicalization: Evidence from the History of a Composite Predicate Construction in English. In: Elizabeth C. Traugott & Bernd Kortmann (eds.), *Constructional Approaches to English Grammar*, 33-67. Berlin/New York: Mouton De Gruyter.
Trousdale, Graeme. 2010. Issues in Constructional Approach to Grammaticalization in English. In: Katerina Stathi, Elke Gehweiler & Ekkehard König (eds.), *Grammaticalization: Current Views and Issues* (Studies in Language Companion Series), 51-72. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Trousdale, Graeme. 2012. Grammaticalization, Constructions and the Grammaticalization of Constructions. In: Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and Language Change: New Reflections*, 167-198. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Trousdale, Graeme. 2013. Multiple Inheritance and Constructional Change. *Studies in Language* 37 (3), 491-514.
Trousdale, Graeme. 2014. On the Relationship between Grammaticalization and Constructionalization [11]. *Folia Linguistica* 48 (2), 557-578.
Trousdale, Graeme & Nikolas Gisborne. 2008. *Constructional Approaches to English Grammar*. Berlin/New York: Mouton De Gruyter.
Tuggy, David. 1993. Ambiguity, Polysemy and Vagueness. *Cognitive Linguistics* 4 (3), 273-290.
Tummers, José, Dirk Speelman & Dirk Geeraerts. 2005. Inflectional Variations in Belgian and Netherlandic Dutch. In: Nicole Delbecque, Johan Van der Auwera & Dirk Geeraerts (eds.), *Perspectives on Variation: Sociolinguistic, Historical, Comparative*, 93-110. Berlin/New York: Mouton De Gruyter.
Van Beveren, Amélie. 2015. *De Intensiverende Resultatiefconstructie. Een Contrastief Corpusonderzoek Nederlands – Afrikaans*. Bachelor dissertation, Universiteit Gent.
Van de Velde, Freek. 2011. Left-Peripheral Expansion of the English Np. *English Language and Linguistics* 15 (2), 387-415.
Van de Velde, Freek. 2014. Degeneracy: The Maintenace of Constructional Networks. In: Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the Scope of Construction Grammar*, 141-180. Berlin/New York: Mouton De Gruyter.
Van den Toorn, Maarten C. 1997. *Geschiedenis Van De Nederlandse Taal*. Amsterdam: Amsterdam University Press.
Van der Horst, Joop & Fred Marschall. 2000. *Korte Geschiedenis Van De Nederlandse Taal*. Den Haag: SDU Uitgevers.
Van der Wouden, Ton & Ad Foolen. 2017. A Most Serious and Extraordinary Problem. Intensification of Adjectives in Dutch, German and English. *Leuvense Bijdragen* 101, 82-100.
Van Krieken, Kobie & José Sanders. 2016a. Diachronic Changes in Forms and Functions of Reported Discourse in News Narratives. *Journal of Pragmatics* 91, 45-59.
Van Krieken, Kobie & José Sanders. 2016b. Framing Narrative Journalism as a New Genre: A Case Study of the Netherlands. *Journalism*,
Van Krieken, Kobie & José Sanders. 2016c. Gruwelijke Moorden, Meeslepende Verhalen: Gebruik En Ontwikkeling Van Narratieve Technieken in Journalistieke Teksten (1850–1939). In: D. Van de Mieroop, L. Buysse & R. Coesemans (eds.), *De Macht Van De Taal: Taalbeheersingsonderzoek in Nederland En Vlaanderen*, 117-130. Leuven: ACCO.
van Trijp, Remi & Luc Steels. 2012. Multilevel Alignment Maintains Language Systematicity. *Advances in Complex Systems* 15 (3), 1-30.
Vanden Wyngaerd, Guido J. 2001. Measuring Events. *Language* 77 (1), 61-90.
Vandewinkel, Sigi & Kristin Davidse. 2008. The Interlocking Paths of Development to Emphasizer *Pure*. *Journal of Historical Pragmatics* 9 (2), 255-287.

Verhagen, Arie. 2002. From Parts to Wholes and Back Again. *Cognitive Linguistics* 13 (4), 403-439.
Verhagen, Arie. 2003a. *The Dutch Way* (Usage-Based Approaches to Dutch. Lexicon, Grammar, Discourse). Utrecht: LOT.
Verhagen, Arie. 2003b. Hoe Het Nederlands Zich Een Eigen Weg Baant. Vergelijkende En Historische Observaties Vanuit Een Constructie-Perspectief. *Nederlandse Taalkunde* 8, 328-346.
Vis, Kirsten. 2011. *Subjectivity in News Discourse. A Corpus Linguistic Analysis of Informalization*. PhD dissertation, Vrije Universiteit Amsterdam.
Vis, Kirsten, José Sanders & Wilbert Spooren. 2009. Subjectiviteit Door De Jaren Heen: Conversationalisatie in Journalistieke Teksten. In: Wilbert Spooren, Margreet Onrust & José Sanders (eds.), *Studies in Taalbeheersing 3*, 405-418. Assen: Van Gorcum.
Vis, Kirsten, José Sanders & Wilbert Spooren. 2012. Diachronic Changes in Subjectivity and Stance: A Corpus Linguistic Study of Dutch News Texts [6//]. *Discourse, Context & Media* 1 (2–3), 95-102.
von Mengden, Ferdinand & Evie Coussé. 2014. The Role of Change in Usage-Based Conceptions of Language. In: Evie Coussé & Ferdinand von Mengden (eds.), *Usage-Based Approaches to Language Change*, 1-19. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Waksler, Rachelle. 2012. *Super*, *Uber*, *So*, and *Totally*: Over-the-Top Intensification to Mark Subjectivity in Colloquial Discourse. In: Nicole Baumgarten, Inke Du Bois & Juliane House (eds.), *Subjectivity in Language and Discourse*, 17-32. Bingley: Emerald Group Publishing Limited.
Washio, Ryuichi. 1997. Resultatives, Compositionality and Language Variation. *Journal of East Asian Linguistics* 6, 1-49.
Wechsler, Stephen & Bokyung Noh. 2001. Predication and Anaphora: Parallels between Korean and English Resultatives. *Language Sciences* 23, 391-423.
Westin, Ingrid & Christer Geisler. 2002. A Multi-Dimensional Study of Diachronic Variation in British Newspaper Editorials. *ICAME Journal* 26, 133-152.
Wharton, Tim. 2016. That Bloody So-and-So Has Retired: Expressives Revisited. *Lingua* 175-176, 20-36.
Wolfe, Tom. 1972. The Birth of 'the New Journalism'; Eyewitness Report. In *New York Magazine*. Accessed online at http://nymag.com/news/media/47353/ on August 16th 2017.
Wulff, Stefanie, Anatol Stefanowitsch & Stefan Thomas Gries. 2007. Brutal Brits and Persuasive Americans: Variety-Specific Meaning Construction in the *into*-Causative. In: Günter Radden, Klaus-Michael Köpcke, Thomas Berg & Peter Siemund (eds.), *Aspects of Meaning Construction in Lexicon and Grammar*, 265-281. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Zeldes, Amir. 2012. *Productivity in Argument Selection from Morphology to Syntax*. Berlin/New York: Mouton De Gruyter.
Zenner, Eline, Jocelyn Daems & Dirk Geeraerts. 2013. Pluricentric Languages and Exocentric Norms. Lexical Convergence between Belgian Dutch and Netherlandic Dutch. Paper presented at Societas Linguistica Europaea (SLE), Split.
Zeschel, Arne. 2012. *Incipient Productivity: A Construction-Based Approach to Linguistic Creativity*. Berlin/New York: Mouton De Gruyter.
Zwicky, Arnold M. & Jerrold M. Sadock. 1975. Ambiguity Tests and How to Fail Them. In: John P. Kimball (ed.), *Syntax and Semantics 4*, 1-36. New York: Academic Press.

Appendices

Data sets

The data sets that were used for the analyses presented in this thesis are available upon request.

Appendix III-1 Synchronic data set, round 1, input intensifiers

5 hoedjes
barstensvol
belazerd
beroerd
bewusteloos
blauw
blind
de bagger
de ballen (uit de broek)
de benen uit het lijf
de beroerte
de bibbers
de blaren
de blubber
de breuk
de buik krom
de hell [sic]
de joeperdepoep
de klaplazerus/klaplazarus
de klere/kolere
de knetters
de knopen van de broek
de krampen
de lazarus/lazerus
de longen uit het lijf
de mikmak
de naad uit het lijf
de neten
de ogen uit (het hoofd)
de oren van het hoofd
de pelikanenschurft
de pest
de pestpleuris/pestpleures
de pestpokken
de pieperdepiep
de pispleuris/pispleures
de pletter
de pleures/pleuris
de pokken
de rambam
de schompus/schompes
de schrikschrak
de shit
de stuipen
de tandjes
de tering
de tierelier
de tieten van het lijf
de touwtjes
de tranen
de tyfus/typhus
de vanalles en nog wat
de vellen
de vinketering
de wimwam
de ziekte
de ziel uit het lijf
de zolen van de schoenen
dood
doof
drie slagen in de rondte
een aap
een apenstaartje
een beroerte
een biet
een bommahoedje
een breuk
een bult
een coma
een deuk
een ei lek
een gat in de lucht
een gil
een glosterhoedje
een gluut
een hartinfarct
een hartstilstand
een hartverknettering
een hartverlamming
een hartvervetting
een hartverzakking

een hoed
een hoedje
een hoedjesverzakking
een hoofddoekje
een indigestie
een kankerhoedje
een kriek
een ongeluk
een pizza
een pleuris
een puntmutsje
een rendierhoedje
een roes
een rolberoerte
een rotje
een Sableye
een sjaaltje
een slag in de rondte
een stuk in de kraag
een stuip
een tulbandje
een veertje
een verlepping
een voetje
een (zwerver) petje
een ziekte
failliet
geel
gek
groen (en geel)
halfdood
het apelazerus/apelazarus
het apezuur/apenzuur
het habbiebabbie/
habbieboebie
het hellybelly
het hoedje
het konijnenzout
het lamlazerus/lamlazarus
het lazarus/lazerus
het lebbes
het laplazerus/laplazarus
het leplazerus/leplazarus
het ongans
het ongeluk
het ongelukje
het pleuris
het rambam
het schompes/schompus
het snot voor de ogen
het vel van de vingers
het vuur uit de sloffen
het wigwam
het zuur
in het zweet
kapot
klem
kreupel
krom
lam
laveloos
lens
levenloos
mal
misselijk
ongans
ongelukkig
plat
purper
raar
rond
rot
scheel
scheurbuik
schor
slap
steendood
stijf
stuk
suf
te barsten
te pletter
tot barstens toe
tot huilen aan toe
tranen
uit de naad
verrot
wezenloos
wild
ziek

Appendix III-2: Synchronic data set, all intensifiers types[79]

	Netherl. Dutch	Belgian Dutch
[n°] slagen in de rondte	9	0
adellijk blauw	0	1
apelazerus	0	1
beten en scheten	0	1
bewusteloos	1	0
bicblauw	0	1
blauw	28	307
blauw en paars	0	1
blind	3	0
bont en blauw	1	1
de balg uit het lijf	0	1
de benen PREP het gat	1	0
de benen PREP het lijf	11	25
de blaren	0	2
de blaren op de tong	8	1
de blaren op de voeten	1	0
de blaren op het verhemelte	1	0
de blubber	0	1
de kleren van het lijf	0	1
de kolere	2	0
de longen uit het lijf	12	38
de naad uit de broek	0	6
de naad uit het lijf	0	12
de nieren los	0	1
de ogen uit het hoofd	6	7
de oren van het hoofd	0	1
de pest	1	0
de pleuris	3	39
de pleuris uit het lijf	0	1
de poten van onder de keukenstoel	0	1
de schoenen vanonder hun voeten	0	1
de tering	4	1
de typhus	0	1
de vingers blauw	1	0
de voeten van onder het lijf	0	4
de ziel uit de naad	0	1
de ziel uit de raap	0	1
de ziel uit het lijf	0	98
de zolen van de schoenen	0	2
donkerblauw	0	1
dood	162	237
een aap	0	33
een beroerte	0	13
een breuk	7	26
een bult	1	53
een delirium	3	4
een delirium tremens	0	1
een deuk	1	1
een eind in de rondte	2	0
een hartaanval	1	1
een hoedje	28	200
een houten hart	1	0
een kontzweer	0	1
een kriek	5	11
een liesbreuk	0	1
een ongeluk	60	58
een ootje	1	0
een pissebed	0	1
een punthoofd	0	5
een rolberoerte	1	0
een rotje	1	0
een slag in de rondte	37	0
een stuip	0	1
een stuk in de gilet	0	1

[79] The unmarked intensifiers were retrieved in round one, the intensifiers that are marked in bold were added in the second search round.

een stuk in de kraag	4	13
een stuk in de voeten	0	3
een zoeavenmuts	1	0
gaar	0	1
gek	10	10
groen	0	2
groen en geel	44	3
halfdood	2	2
het apelazerus	6	0
het apenzweet	1	0
het apezuur	1	3
het geel	0	1
het hart uit het lijf	0	4
het lazerus	4	1
het leplazerus	6	1
het licht uit	0	1
het schompes	6	0
het snot voor de ogen	4	1
het vel van de botten	1	0
het vuur uit de schoenen	0	1
het vuur uit de sloefen	0	2
het vuur uit de sloffen	14	24
het vuur uit de slofkens	0	1
het vuur uit de sokken	1	0
het zwart voor de ogen	0	1
in het zweet	34	93
kapot	117	125
klem	15	38
krankjorum	1	1
kreupel	0	3
krom	3	28
lam	7	7
murw	0	9
ongans	11	0
onnozel	0	3
paars	1	1
plat	1	1
pleuris	0	2
rot	154	241
scheel	5	0
schor	6	24
slap	3	1
spinaziegroen	0	1
steendood	0	16
stuk	4	4
suf	137	139
te barsten	8	6
te pletter	93	550
te pleuris	0	6
te sappel	0	1
uit de naad	21	192
verloren	0	1
verrot	1	5
VLD-blauw	0	1
wezenloos	22	4
wild	34	0
ziek	3	19
zot	0	13
	1190	2818

Appendix III-3: Synchronic data set, all verb types[80]

	Netherl. Dutch	Belgian Dutch
aaien	1	0
adverteren	3	1
analyseren	0	2
applaudisseren	0	2
associëren	1	0
babbelen	0	1
bellen	8	6
beminnen	1	0
besparen	0	1
betalen	18	167
betogen	0	1
bezetten	0	1
bezuinigen	1	0
bibberen	0	1
bidden	1	0
blaffen	0	2
blazen	2	4
blokken	0	1
blowen	3	0
boetseren	1	0
borduren	1	0
breien	0	1
brullen	0	2
bula'en [sic]	0	1
chatten	0	1
cijferen	0	1
communiceren	2	2
concurreren	0	1
consumeren	2	3
dansen	5	15
debatteren	2	0
demarreren	0	1
denken	2	9
dirigeren	1	0
discussiëren	2	2
doperen	0	1
downloaden	1	0
draven	0	2
drinken	26	53
dromen	1	0
dubben	0	1
duiken	1	0
duwen	1	0
e-mailen	1	0
eten	18	15
experimenteren	1	0
feesten	1	0
fietsen	4	20
fluiten	0	1
foeteren	0	1
fuiven	0	1
gamen	0	1
gebruiken	1	0
genieten	1	1
gillen	1	2
gluren	1	0
gniffelen	0	1
gokken	1	0
golfen	1	0
googlen	0	2
grabbelen	0	1
grijnzen	1	0
groeien	1	1
handelen	1	0
hoesten	0	2
hollen	1	4
hongeren	2	2
huilen	0	6
huren	0	1
ideologiseren	1	0
internetten	1	1

[80] The unmarked verbs were retrieved in round one, the verbs that are marked in bold were added in the third and final search round.

investeren	0	1
isoleren	0	1
jagen	0	1
janken	0	1
jazzen	0	1
joggen	0	1
juichen	0	2
kakelen	0	1
kakken	0	1
kandideren	1	0
kiezen	1	0
kijken	0	6
klagen	1	0
kletsen	3	0
klikken	4	0
kloppen	0	2
knagen	1	0
knokken	4	2
knuffelen	0	1
koersen	0	2
koken	0	1
kopen	8	7
kotsen	1	0
krabben	1	1
kreunen	0	1
lachen	70	123
lenen	1	1
leren	4	0
leuteren	0	1
lezen	6	5
liegen	0	1
lijnen	1	1
lobbyen	1	1
lopen	42	151
lullen	1	1
mailen	1	2
manipuleren	1	0
manoeuvreren	1	0
mediteren	0	1
meppen	0	1
musterberen [sic]	0	1
neuken	6	1
nuanceren	0	1
oefenen	0	4

onderhandelen	2	3
organiseren	2	0
orkestreren	0	1
overleggen	2	0
paaien	1	0
paffen	1	2
patrouilleren	0	1
peilen	0	1
peinzen	8	7
pendelen	0	1
piekeren	22	68
pingelen	1	0
plannen	0	1
ploeteren	0	1
poetsen	0	1
prakkiseren	1	0
praten	9	3
prikken	0	1
printen	1	0
procederen	1	0
produceren	0	1
programmeren	0	1
protesteren	1	0
prutsen	0	1
puzzelen	0	1
raden	1	0
rappen	0	2
ravotten	0	1
recyclen	1	0
redeneren	0	1
regeren	0	3
registreren	1	0
reizen	2	2
rekenen	2	3
relativeren	1	2
rennen	11	17
repeteren	0	5
rijden	22	234
roepen	0	4
roken	6	5
sakkeren	0	1
schieten	7	6
schijnen	0	1
schilderen	1	0

schitteren	0	1
schnabbelen	1	0
schreeuwen	3	38
schrijven	5	3
schrikken	223	467
scrabbelen	0	1
selecteren	1	0
serveren	1	0
shoppen	2	5
signalen [sic]	0	1
sjouwen	2	0
slaan	1	0
slapen	1	0
slempen	1	0
slepen	1	0
sleuren	1	0
sleutelen	0	1
slikken	0	3
smokkelen	1	0
sms'en	4	1
snoepen	0	2
snoeven	0	1
snuiven	4	1
solliciteren	4	2
sparen	1	4
spelen	2	26
speuren	1	0
sponsoren	0	2
sporten	0	6
springen	0	1
spuiten	2	1
spurten	0	4
staken	0	2
stampen	0	2
staren	0	2
steken	1	1
stoken	1	0
studeren	1	4
supporteren	0	2
surfen	1	1
swingen	1	0
tappen	0	1
tekenen	1	0
telefoneren	1	2
tennissen	1	2
tikken	1	0
tillen	2	1
time-managen	0	1
tobben	0	2
toeren	1	0
toeteren	1	0
tongzoenen	1	0
trainen	13	22
transformeren	1	0
trappen	3	4
treuren	0	1
turen	0	1
turnen	0	1
turven	1	0
uitleggen	0	1
vallen	0	1
varen	0	1
vechten	11	9
vegen	1	0
vergaderen	4	6
vergelijken	0	1
verkopen	1	0
verschieten	0	6
vliegen	18	11
vloeken	0	3
vragen	0	1
vreten	3	6
vrijen	1	1
waarschuwen	0	1
wassen	0	1
werken	135	402
werven	1	0
wiebelen	0	1
winkelen	0	1
wroeten	1	8
zappen	1	0
zeulen	0	2
zeuren	0	1
zich amuseren	6	181
zich enerveren	0	1
zich ergeren	133	232
zich generen	3	7
zich integreren	1	0

zich isoleren	0	1
zich schamen	80	43
zich verheugen	1	0
zich vermaken	1	0
zich vervelen	58	80
zingen	5	13
zoeken	15	44

zuigen	1	0
zuipen	16	31
zwemmen	4	0
zweten	1	28
zwijgen	1	2
zwoegen	0	8
	1190	2818

Appendix III-4 Diachronic data set, round 1, output intensifiers

arm
beroerd
bewusteloos
blaren
blauw(er)
blauw en groen
bleek
blind
bloot
bont en blauw
buikkrampen
buikpijn
de benen uit het gat
de benen uit het lijf
de blubber(s)
de griebels
de hik
de kelen schor
de klere/kolere
de ledematen uit de gewrichten
de longen uit het lijf
de mazelen
de naad uit het lijf
de ogen uit (het hoofd)
de pest
de pestpokken
de pleuris/pleures/pleurus
de pokken
de poten stuk
de poten uit het lijf
de rambam
de stuipen
de stuipjes
de takken
de tering
de tering-takke
de tranen
de vingers krom
de vingers ten bloede
de vinketering
de voeten stuk
de zenuwen
de zolen van de voeten
dood
een aanp [sic]
een aap(je)
een barst
een (halve) beroerte
een blauw hart
een breuk
een buil
een bult
een deuk
een (gevaarlijke) ziekte
een hartinfarct
een hartverlamming
een hoed(je)
een kokosnoot
een (onfatsoenlijke) koliek
een kriek
een loei
een mik
een ongeluk
een petje
een puist
een pukkel
een punthoofd
een rotberoerte
een rotje
een slaghoedje
een slag in de rondte
een stuip
een zuurstok
flauw
gaar
geel
geel en groen
gek
gezond
grijs
groen
groen en blauw
groen en geel
halfdood
halfgek
halfkapot
halflam
halfziek
het apelazerus
het apezuur/apenzuur
het bloed in de schoenen
het habbi-babbi
het hoedje
het hoofd suf
het laplazerus/laplazarus
het lazerus/lazarus
het leplazerus/leplazarus
het licht uit de ogen

het ongeluk	lam	stuk
het pleuris	laveloos	suf
het rambam	lazerus/lazarus	te barsten
het schompes/schompus	leeg	te blubber
het schuim op de hiel	lens	te pletter
het schuim op de ziel	[n°] slagen in de rondte	ten doode
het snot voor de ogen	ondersteboven	ten doode toe
het vuur uit de molières	ongans	tranen (in de ogen)
het vuur uit de schoenen	ongelukkig	tureluurs
het vuur uit de sloffen	paars en groen	uit de naad
het vuur uit de spikes	rond	uit de naden
het zuur	rood	uit de sloffen
in de kreukels	rood en groen	uit het lid
in de poeier	rot	uit het lood
in het zweet	scheef	verrot
in pust	scheel	wezenloos
kapot	schor	wild
klem	slap	witjes
kleurenblind	steendood	zenuwziek
krampen	stom	ziek
krom	stuipen	zwart

Appendix IV-1: Verbs in the intensifying subset of SoNaR

	Netherl. Dutch	Belgian Dutch
aaien	1	0
adverteren	3	1
analyseren	0	2
applaudisseren	0	2
associëren	1	0
babbelen	0	1
bellen	8	6
besparen	0	1
betalen	18	167
betogen	0	1
bezetten	0	1
bezuinigen	1	0
bibberen	0	1
bidden	1	0
blaffen	0	2
blazen	2	4
blokken	0	1
boetseren	1	0
borduren	1	0
breien	0	1
brullen	0	1
bula'en [sic]	0	1
chatten	0	1
cijferen	0	1
communiceren	2	2
consumeren	2	3
dansen	2	11
debatteren	2	0
demarreren	0	1
denken	1	7
dirigeren	1	0
discussiëren	2	2
doperen	0	1
downloaden	1	0
draven	0	1
drinken	11	31

dromen	1	0
dubben	0	1
duiken	1	0
duwen	1	0
e-mailen	1	0
eten	12	7
experimenteren	1	0
feesten	1	0
fietsen	3	17
foeteren	0	1
fuiven	0	1
gamen	0	1
genieten	1	1
gillen	1	2
gluren	1	0
gniffelen	0	1
gokken	1	0
golfen	1	0
googlen	0	2
grabbelen	0	1
grijnzen	1	0
handelen	1	0
hoesten	0	2
hollen	1	4
huilen	0	6
huren	0	1
ideologiseren	1	0
internetten	1	1
investeren	0	1
janken	0	1
jazzen	0	1
joggen	0	1
juichen	0	2
kakelen	0	1
kakken	0	1
kandideren	1	0
kiezen	1	0
kijken	0	6
klagen	1	0
kletsen	3	0
klikken	4	0
kloppen	0	1
knagen	1	0
knokken	4	2
knuffelen	0	1
koersen	0	1
koken	0	1
kopen	8	7
kotsen	1	0
krabben	1	1
lachen	70	123
lenen	1	1
leren	4	0
leuteren	0	1
lezen	5	4
liegen	0	1
lijnen	1	1
lobbyen	1	1
lopen	36	125
lullen	1	1
mailen	1	2
manipuleren	1	0
mediteren	0	1
meppen	0	1
musterberen	0	1
neuken	5	1
nuanceren	0	1
oefenen	0	4
onderhandelen	2	3
organiseren	2	0
orkestreren	0	1
overleggen	2	0
paaien	1	0
patrouilleren	0	1
peilen	0	1
peinzen	8	6
pendelen	0	1
piekeren	22	65
pingelen	1	0
plannen	0	1
ploeteren	0	1
poetsen	0	1
prakkiseren	1	0
praten	9	2
printen	1	0
procederen	1	0
produceren	0	1
programmeren	0	1

protesteren	1	0
prutsen	0	1
puzzelen	0	1
raden	1	0
rappen	0	2
ravotten	0	1
recyclen	1	0
redeneren	0	1
registreren	1	0
reizen	2	2
rekenen	1	3
relativeren	1	2
rennen	9	17
repeteren	0	5
rijden	4	69
roepen	0	4
roken	4	4
sakkeren	0	1
schijnen	0	1
schilderen	1	0
schitteren	0	1
schnabbelen	1	0
schreeuwen	1	29
schrijven	5	3
schrikken	223	465
scrabbelen	0	1
serveren	1	0
shoppen	2	4
signalen	0	1
sjouwen	2	0
slempen	1	0
slepen	1	0
sleuren	1	0
sleutelen	0	1
slikken	0	1
smokkelen	1	0
sms'en	4	1
snoepen	0	2
snoeven	0	1
snuiven	2	1
solliciteren	4	2
sparen	1	4
spelen	0	22
speuren	1	0
sponsoren	0	2
sporten	0	4
springen	0	1
spurten	0	4
staken	0	2
stampen	0	2
staren	0	2
stoken	1	0
studeren	1	4
supporteren	0	2
surfen	1	1
swingen	1	0
tappen	0	1
tekenen	1	0
telefoneren	1	2
tennissen	1	1
tikken	1	0
tillen	1	1
time-managen	0	1
tobben	0	2
toeren	1	0
toeteren	1	0
tongzoenen	1	0
trainen	11	21
transformeren	1	0
trappen	2	2
treuren	0	1
turen	0	1
turnen	0	1
turven	1	0
uitleggen	0	1
vechten	3	6
vegen	1	0
vergaderen	4	6
vergelijken	0	1
verkopen	1	0
verschieten	0	6
vloeken	0	3
vragen	0	1
vreten	1	5
vrijen	1	1
waarschuwen	0	1
werken	112	346
werven	1	0

wiebelen	0	1
winkelen	0	1
wroeten	1	7
zappen	1	0
zeulen	0	1
zeuren	0	1
zich amuseren	5	181
zich enerveren	0	1
zich ergeren	133	232
zich generen	3	7
zich integreren	1	0
zich schamen	80	43

zich verheugen	1	0
zich vermaken	1	0
zich vervelen	58	78
zingen	4	12
zoeken	14	44
zuigen	1	0
zuipen	14	25
zwemmen	2	0
zweten	1	28
zwijgen	1	2
zwoegen	0	8
	1042	2445

Appendix IV-2: Intensifiers in the intensifying subset of SoNaR[81]

	Netherl. Dutch	Belgian Dutch
[n°] slagen in de rondte	9	0
adellijk blauw	0	1
apelazerus	0	1
beten en scheten	0	1
bewusteloos	1	0
bicblauw	0	1
blauw	28	307
blauw en paars	0	1
blind	3	0
bont en blauw	1	0
de balg uit het lijf	0	1
de benen PREP het gat	1	0
de benen PREP het lijf	11	25
de blaren	0	1
de blaren op de tong	8	1
de blaren op de voeten	1	0
de blaren op het verhemelte	1	0
de blubber	0	1

de kleren van het lijf	0	1
de kolere	2	0
de longen uit het lijf	12	38
de naad uit de broek	0	6
de naad uit het lijf	0	12
de nieren los	0	1
de ogen uit het hoofd	6	7
de oren van het hoofd	0	1
de pest	1	0
de pleuris	3	39
de pleuris uit het lijf	0	1
de poten van onder de keukenstoel	0	1
de schoenen vanonder hun voeten	0	1
de tering	4	1
de typhus	0	1
de vingers blauw	1	0

[81] The overlapping intensifiers in SoNaR-NL and SoNaR-BE are marked in grey.

de voeten van onder het lijf	0	4
de ziel uit de naad	0	1
de ziel uit de raap	0	1
de ziel uit het lijf	0	98
de zolen van de schoenen	0	2
donkerblauw	0	1
dood	116	177
een aap	0	32
een beroerte	0	13
een breuk	6	26
een bult	1	53
een delirium	3	4
een delirium tremens	0	1
een deuk	1	1
een eind in de rondte	2	0
een hartaanval	0	1
een hoedje	28	200
een houten hart	1	0
een kontzweer	0	1
een kriek	5	11
een liesbreuk	0	1
een ongeluk	60	57
een ootje	1	0
een pissebed	0	1
een punthoofd	0	5
een rolberoerte	1	0
een rotje	1	0
een slag in de rondte	37	0
een stuip	0	1
een stuk in de gilet	0	1
een stuk in de kraag	4	13
een stuk in de voeten	0	3
een zoeavenmuts	1	0
gaar	0	1
gek	8	10
groen	0	2
groen en geel	44	3
halfdood	2	2
het apelazerus	6	0
het apenzweet	1	0
het apezuur	1	3
het geel	0	1
het hart uit het lijf	0	4
het lazerus	4	1
het leplazerus	6	1
het licht uit	0	1
het schompes	6	0
het snot voor de ogen	4	1
het vel van de botten	1	0
het vuur uit de schoenen	0	1
het vuur uit de sloefen	0	2
het vuur uit de sloffen	14	24
het vuur uit de slofkens	0	1
het vuur uit de sokken	1	0
het zwart voor de ogen	0	1
in het zweet	10	21
kapot	104	72
klem	9	1
krankjorum	1	1
kreupel	0	3
krom	3	27
lam	7	7
murw	0	3
ongans	11	0
onnozel	0	2
paars	1	1
plat	1	1
pleuris	0	2
rot	154	241
scheel	5	0
schor	3	14
slap	3	1
spinaziegroen	0	1
steendood	0	16
stuk	2	0
suf	123	129
te barsten	8	6
te pletter	60	445

te pleuris	0	6
te sappel	0	1
uit de naad	21	192
verloren	0	1
verrot	0	5
VLD-blauw	0	1

wezenloos	22	2
wild	34	0
ziek	1	11
zot	0	13
	1042	2445

Appendix IV-3: Output covarying collexeme analysis, SoNaR-NL

Available online at https://emmelinegyselinck.wixsite.com/phdthesis2018 or upon request

Appendix IV-4: Output covarying collexeme analysis, SoNaR-BE

Available online at https://emmelinegyselinck.wixsite.com/phdthesis2018 or upon request

Appendix V-1: All verbs per decennium, Delphcorp

All verbs are listed in descending order based on the total of summed frequencies.

	1830-1839	1850-1859	1870-1879	1890-1899	1910-1919	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1989	1990-1995
schrikken					9	47	68	133	103	171	285	470
lachen	1	2	13	12	35	64	74	146	98	117	143	166
zich ergeren		1	15	8	16	18	18	38	33	76	142	270
werken	2	1	7	16	11	19	24	34	46	96	130	224
lopen			2	11	7	8	20	35	26	58	130	134
zich schamen		1	2	6	11	18	22	37	22	34	59	106
zich vervelen			2	3	8	20	20	25	34	34	62	85
piekeren					1	3		5	16	27	39	60
zoeken			2			7	3	9	10	27	18	27
trainen									4	7	25	45
peinzen			1	9	10	4	10	16	8	7	4	9
betalen						2	3	6	8	8	13	37
prakkiseren				2	2	4	8	5	3	10	13	11
sjouwen						2	3	3	6	10	17	14
rijden							1	3	2	9	13	26

denken	1		3	4	10	12	7	8		2		
rennen								1	2	2	10	29
vechten				2	2		2	3	5	10	6	12
schreeuwen		2		2	3	2	1	2	4	5	4	15
schrijven			1	2	2	2	1	5	3	5	7	7
drinken	1				1	2	2	5	5	4	2	12
fietsen							1	3	3	10	3	12
trappen				1		1	3	2	4	11	2	8
zich generen						4	2	4	8	3	4	5
zuipen									2	1	5	20
eten								3	2	2	9	7
kijken				2	2	4		2	2	5	2	2
praten			1			1		3	1	3	4	7
solliciteren								1		1	6	12
klappen			1				1	4	3	1	1	8
knokken										3	5	11
spelen							1	1		2	7	6
zweten			1						3	3	5	5
lezen						1	1	1			7	5
rekenen						1		2	1	1	5	5
kniezen			2		1	6	1	3		1		
bellen										2	4	7
vergaderen								1	2		6	4
vreten								1		3	2	6
blazen			1				1		2	3	1	3
kopen								1	1	2		7
studeren								1	1	1	4	4
adverteren										2	3	5
tillen								2	2	1	3	2
zingen				1						2	1	5
treuren	1			2	3	1	1					
hoesten								3		1	1	2
juichen					2	1		2		1	1	
reizen									1		2	4
draven							3	1		1		1
huilen						2		1			1	2
leren									1	1	3	1
schreien				4	1						1	
vliegen							1		3			2
zich amuseren											4	2
zwoegen					1	2		2				1
dansen						1			1			3
gillen					1			1		1	1	1
oefenen											1	4

repeteren								1	1	2		1
roken										2		3
springen										1	2	2
stoken										2	3	
tellen					1			1		1		2
slaan										1		3
slepen										1	3	
wenen				2	2							
balen											1	2
beuken									1	1	1	
bezuinigen											1	2
blaffen					1					1		1
blokken												3
brullen										1		2
confereren						1					1	1
draaien							1				1	1
hollen										2		1
krabben						1				1		1
naaien									1			2
pennen					1	1				1		
ploeteren					1					1		1
poetsen									1	1	1	
rekken						1	2					
roepen			1	1								1
sparen										1	1	1
staren										1	2	
tobben						3						
zappen												3
zeulen									1		1	1
bladeren												2
bouwen										1		1
discussiëren									1		1	
drukken							1		1			
duiken										2		
fluiten										1		1
fotograferen									1			1
kiezen											2	
kletsen												2
knippen												2
onderhandelen										1		1
regelen								1			1	
roeien								1		1		
sappelen					1						1	
schaatsen											1	1

schakelen										1		1
scheppen								1				1
scheuren												2
schieten												2
sloffen										1	1	
smeren												2
spreken						1					1	
surfen											2	
verzinnen					2							
voetballen										2		
zien				1							1	
zwemmen										2		
acteren											1	
analyseren						1						
annonceren										1		
argumenteren					1							
associëren											1	
baggeren											1	
bakken								1				
bewapenen											1	
bidden										1		
bloeien										1		
boemelen							1					
boenen									1			
borstelen												1
breien											1	
cijferen								1				
citeren												1
combineren										1		
communiceren												1
concurreren												1
congresseren											1	
controleren												1
demarreren											1	
demonstreren												1
dobbelen					1							
dobberen										1		
dreigen										1		
dresseren									1			
duwen								1				
experimenteren												1
feesten							1					
filmen									1			
filosoferen					1							

fingeren									1			
geeuwen										1		
giechelen										1		
gieren										1		
gooien										1		
grappen											1	
graven												1
hijgen						1						
hinniken												1
hongeren												1
investeren											1	
janken						1						
kappen												1
kauwen							1					
klagen												1
klikken												1
knijpen										1		
knipperen												1
kwellen					1							
leggen									1			
lenen											1	
liften										1		
lonken												1
lullen												1
mediteren										1		
mijmeren		1										
molenwieken					1							
neuken												1
organiseren												1
overleggen											1	
pachten				1								
pakken										1		
pezen												1
pijnigen										1		
pleiten												1
pompen					1							
prijzen												1
protesteren											1	
racen										1		
rammelen							1					
recenseren										1		
redeneren							1					
reorganiseren												1
reserveren											1	

roddelen									1			
roeren												1
samplen												1
scanderen											1	
scheiden												1
schetteren							1					
schoppen							1					
schrappen				1								
schreppen [sic]											1	
schuimen												1
schuiven											1	
sikkeneuren									1			
slapen									1			
slempen												1
slikken												1
sloven			1									
snateren							1					
snikken				1								
snoeien												1
snuiten											1	
sprinten												1
spuiten												1
steken											1	
stelen											1	
stomen												1
storen												1
strelen												1
sturen										1		
suffen				1								
swingen												1
telefoneren											1	
toeren												1
tollen									1			
transpireren												1
trommelen												1
trompetteren												1
turen				1								
vallen											1	
verdienen												1
verschieten										1		
verschrikken						1						
verzamelen												1
vloeken										1		
wachten					1							

wankelen											1	
wassen												1
wegen												1
wisselen										1		
worstelen							1					
wrijven								1				
zagen								1				
zich verbazen					1							
zich vermaken												1
zwaaien										1		

Appendix V-2: All intensifiers per decennium, Delphcorp

All intensifiers are listed in descending order based on the total of summed frequencies.

	1830-1839	1850-1859	1870-1879	1890-1899	1910-1919	1930-1939	1940-1949	1950-1959	1960-1969	1970-1979	1980-1989	1990-1995
dood	3	1	26	26	48	111	91	152	131	122	168	243
rot							1	11	29	134	167	257
azuur-blauw											1	
suf	1	1	4	15	22	24	23	33	31	76	88	170
kapot				2	2	2	3	7	38	66	107	137
het vuur uit de sloffen			1	8	7	7	16	28	10	23	100	76
een ongeluk			1	1	7	28	24	36	41	36	31	65
te pletter								2	3	15	87	136
een hoedje							13	37	13	22	49	66
bleek										1		
groen en geel					1	2	2	8	8	24	44	107
bloot										1		
blauw					1	4	7	16	15	19	34	51
de adem uit de longen									1			
uit de naad						1	2		2	19	36	71
wezenloos								1	2	15	46	66
de benen uit de naad											2	
de benen uit het lid						1						
de blaren							1		1			
de blaren in de handen								1				
de blaren op de hakken											1	
een aap					2	8	14	34	17	10	8	8
wild							2	8	5	21	28	33

krom			2	4		13	16	21	7	10	11	12
te barsten								2	7	29	20	27
lam					3	1	3	5	9	13	19	23
de buik rond						1						
halfdood			4	11	17	15	11	7	1	2	3	2
de hakken scheef									1			
de handen blauw								1				
de handen kapot							1	1				
de longen uit het lijf								1	1	8	13	45
de hersenen suf								3	1			
slap				1	5	10	9	11	5	9	9	7
ziek	1	1	8	2	9	5	6	18	9		3	2
het apezuur							2	4	1	6	9	35
een bult		1		1			2	13	2	6	21	10
gek						3	1	10	22	7	6	7
de ledematen uit de gewrichten		1										
de longen leeg							2		1	1		
de longen stuk				1								
de longen te barsten		1										
de ogen uit het hoofd		1	1	6	8	7	6	3		2	8	13
in het zweet				2	2	1	3	9	3	6	11	17
de mazelen								1				
een kriek						2	10	16	4		5	13
de nagels blauw									1			
de nekspieren uit het lid							1					
de ogen blind				2								
de ogen uit							1					
de ogen uit de kassen										1		
de benen PREP het lijf								4	6	9	12	18
de ogen zat								1				
de oren rood											1	
bont en blauw					1				2	3	10	32
een breuk								3	3	3	7	19
schor		1	2	2	6	6	3	4	4	1	2	4
de pleuris										2	8	23
de poten kapot										2		
het lazerus									2	8	8	15
een stuk in de kraag					1	1	2	5	4	3	3	9
de stuipen								1		3		
de stuipjes										1		
ongelukkig							1	5	8	7		2
het hoofd suf					2		1	2	2	3	5	7
een rotje							1	1	9	5	1	4
de tranen								1				

scheel									1		4	15
een beroerte					1		2	5	4	4	1	2
de vingers groezelig					1							
de vingers krom			2	1				1		1		
de vingers moe				1								
de vingers wond en rond									1			
groen							2	5		4	5	3
het vuur uit de schoenen				2			2	4	1	4	3	3
het leplazerus											6	10
de ziel dood			1									
de zolen PREP de voeten										1		
de zolen uit de sloffen				1								
ongans									1	2	4	9
een aanp								1				
het apelazerus									1	1	4	9
een aapje							1					
een barst										1		
over de kop							1	3		3		8
de blubber									1	1	6	6
de vingers blauw					2	3		1	2	2	1	3
de handen stuk			1					1	3	1	1	6
klem										1	2	9
een halve beroerte								1				
een hart in het lijf	1											
een hartverlamming						1						
het licht uit de ogen										5	2	4
een deuk											2	8
een kokosnoot								1				
een koliek					1	1		1				
groen en blauw							1				8	1
een loei										1		
een mik						1						
uit de naden											6	3
de blaren op de tong										2	1	5
de pokken								1	1	2		4
een punthoofd								1				
een rotberoerte										1		
het laplazerus										3	3	2
leeg							3			4		1
een slaghoedje									1	1		
een stuip			1		1		2	1		1		
een stuk in de hakken											1	
beroerd									2		2	3
de pestpokken											2	5

een zuurstok								1				
flauw					1							
de zenuwen									2	4		1
geel										1		
het schompes										1	2	4
een pukkel						1	1			2	1	1
grasgroen								1				
een slag in de rondte										1	1	4
grijs en groen											1	
geel en groen							1	1		2		2
verrot										1	1	4
blind			1	2	1							1
groen en grijs							1					
kleurenblind						1	2			1		1
halfgek				1		1						
halfkapot										1		
halflam				1		1	2			1		
halfslap						1						
halfsuf					1	2				2		
halfziek						1						
lens								1			1	3
stuk							1			1	1	2
het hoedje								1				
de klere										1	1	2
de kolere										2	1	1
een hoed						1		1	1			1
grijs									3			1
het ongans											3	1
het snot voor de ogen												4
de benen PREP het gat										1		2
de blaren PREP de voeten											2	1
de naad uit het lijf										1		2
het schuim op de hiel										1		
het schuim op de mond								1		1		
het schuim op de ziel										1		
de pest							1					2
de poten PREP het lijf									1	1		1
het vuur uit de pen								1				
het vuur uit de rennerssloffen								1				
het vuur uit de schaatsen									1			
het vuur uit de schenen							1		2	2	1	
de rambam									1			2
rond								1				2
het vuur uit de slofjes									1			

het vuur uit de spaken											1	
het vuur uit de spikes										1		
het vuur uit de sportsloffen										1		
het vuur uit de vingers								1				
het vuur uit de voetbalschoenen							3					
te blubber											1	2
in de poeier										1		
[n°] slagen in de rondte												2
bewusteloos							1					1
de krampen									1			1
de takken												2
de vinketering												2
krampen										1		
een puist										1		1
gaar												2
het rambam								1				1
lazerus										1		
het zuur								1				1
laveloos												2
apelazerus												1
blaren												1
blaren op de tong												1
blauw en groen												1
paars									1		1	
de blaren op de tond												1
de bril van het hoofd												1
de griebels												1
rood, wit en blauw							1					
de heupen stuk												1
scheef										1		
de hik												1
de longen uit de balg												1
de tering												1
stom									1	1		
stuipen					1		1					
de tering-takke												1
de vingers beurs												1
de voeten PREP het lijf												1
een blauw hart												1
een stuk in de kont												1
ten doode			1	3	1							
tranen					1	1	2	5	2		2	
tranen met tuiten							2					

tureluurs										2		
het hoofd gek												1
het vuur uit de molières												1
uit het lid							1	1				
uit het lood										2		
in pust												1
ledematen blauw												1
paars en groen												1
zenuwziek								1				
rood en groen												1
ziek en weer gezond						1						
zwart						1				1		

Appendix V-3: Output covarying collexeme analysis, Delphcorp

Available online at https://emmelinegyselinck.wixsite.com/phdthesis2018 or upon request

Samenvatting

Theoretische context en methodologie

Deze thesis presenteert de resultaten van een synchroon en diachroon corpusonderzoek naar de rol van productiviteit bij de (re)organisatie van het constructionele netwerk. Hierbij richten we ons eerst op de factoren die van belang zijn bij het bepalen van de productiviteit van een constructie en hoe we de impact van die factoren kunnen meten. Op basis daarvan wordt onderzocht in welke mate productiviteit (op zowel het hoogste niveau als het laagste niveau, d.w.z. analogie) een rol heeft gespeeld bij de organisatie van en interne verschuivingen binnen het constructionele netwerk. Hiervoor focussen we op een specifieke constructie, de intensiverende pseudoreflexieve resultatiefconstructie, geïllustreerd in volgende voorbeelden:

(285) Overtreders, mensen die beboetbaar bezig zijn geweest, schrikken zich vaak een hoedje over de hoogte van de boete. (SoNaR)

(286) Ik zapte er nu langs en schrok me de blaren van het geplamuurde gezicht met die rode lippen (Twitter, 23/10/2016)

Die constructie vertoont in het hedendaags Nederlands een interessante mix van productiviteit en lexicale idiosyncrasie. Hoewel de constructie aan de ene kant heel wat ruimte biedt voor talige creativiteit, zie (286), zijn er aan de andere kant toch een aantal "vaste combinaties" of collocationele voorkeuren die de creativiteit enigszins binnen de perken houden, zie (285). In deze thesis volgen we de recente geschiedenis van de constructie en gaan we na welke constructionele veranderingen ze zoal heeft ondergaan sinds het begin van de 19de eeuw. Gezien haar intensiverende betekeniscomponent laat de constructie ons tevens toe om een blik te werpen op de mogelijke invloed van expressiviteit op de veranderingen die de constructie heeft ondergaan. Uit verschillende studies blijkt namelijk dat taalgebruikers, zeker binnen het domein van intensivering, een zekere universele drang naar expressiviteit voelen, in die zin dat ze zich door middel van hun taal (bewust of onbewust) willen onderscheiden van hun medetaalgebruiker. In hedendaags taalgebruik komt dat tot uiting in een zekere zin voor talige creativiteit, maar ook vanuit diachroon perspectief is er een belangrijke rol weggelegd voor die expressiviteit. Wanneer een bepaalde uitdrukking te frequent wordt en/of wijd verspreid geraakt in de talige gemeenschap, kan die na verloop van tijd haar expressieve kracht verliezen. De uitdrukking kan dan toe zijn aan vervanging – of tenminste aanvulling – door een nieuwe,

expressievere variant. Dat zorgt ervoor dat er constant nieuwe vormen worden geïntroduceerd (Stoffel 1901, Bolinger 1972, Partington 1993, Lorenz 2002, De Clerck & Colleman 2013).

Concreet zijn we in een omvangrijk corpus van hedendaags en historisch krantenmateriaal, samengesteld op basis van materiaal uit Delpher (voor de periode 1830-1995, enkel Nederlands Nederlands) en SoNaR (vanaf 1995, zowel Belgisch als Nederlands Nederlands) op zoek gegaan naar een zo exhaustief mogelijke set aan voorbeelden van de intensiverende pseudoreflexieve resultatiefconstructie. We hebben dit gedaan aan de hand van een cyclische zoekmethode, waarbij de output van de ene zoekopdracht opnieuw werd gebruikt als input voor de volgende zoekopdracht. Na meerdere zoekrondes en een uitgebreid proces van manuele selectie hebben we een dataset samengesteld van 3487 voorbeelden voor het hedendaags Nederlands (1042 voor Nederlands Nederlands en 2445 voor Belgisch Nederlands) en 6137 voor de periode tussen 1830 en 1995. De volledige dataset werd vervolgens geannoteerd voor verschillende variabelen, o.a. lemma werkwoord, reflexiviteit en transitiviteit van het werkwoord, lemma intensiveerder, syntactische categorie van de intensiveerder, vorm van het reflexief voornaamwoord en nationale variëteit als enige extra-linguïstische variabele.

Resultaten

1. Algemene frequentie en gebruik. Een eerste blik op de synchrone data toont dat de constructie in het hedendaags Nederlands een grote variatie vertoont, met in totaal 260 verschillende werkwoorden en 122 verschillende intensiveerders in de volledige SoNaR dataset. Het feit dat zowat de helft van die werkwoordstypes en meer dan een derde van de intensiveerdertypes slechts één keer voorkomen (dit zijn de hapax legomena), bevestigt het creatieve potentieel van de constructie. Tegelijk is het zo dat bepaalde individuele werkwoorden en intensiveerders (en specifieke werkwoord-intensiveerdercombinaties, cf. infra) veel frequenter voorkomen dan andere, wat toont dat er ook een grote mate van conventionaliteit meespeelt. Semantisch gezien kunnen de werkwoorden allerlei activiteiten uitdrukken waarvan een of ander inherent aspect geïntensiveerd kan worden – al is er een zekere voorkeur voor werkwoorden die een emotionele gewaarwording of fysiek belastende handeling uitdrukken. De intensiveerders kunnen verschillende syntactische vormen aannemen, vb. AC, NC, VzC, NC+VzC en NC+AC, en vertonen ook een zekere semantische diversiteit. Een groot deel van de meest frequent gebruikte intensiveerders drukken een negatieve toestand uit (vb. *dood, rot, suf, te pletter*, *uit de naad*). Daarnaast zijn er ook enkele intensiveerders waarin een onvervreemdbaar lichaamsdeel of kledingstuk centraal staat (vb. *de longen uit het lijf, de ziel uit het lijf* of *het vuur uit de sloffen),* een opvallend gevarieerde groep van ziektetermen (vb. *de pleuris, de tering, de tyfus, het apelazerus, het leplazerus, het schompes*) en enkele kleurtermen (vb. *blauw, groen en geel, groen*). Met uitzondering van die laatste categorie lijkt het erop dat intensiveerders vooral gerekruteerd worden uit domeinen waar een negatieve connotatie aan vasthangt. Tot slot zijn er nog enkele losse intensiveerders die moeilijk in één van de eerdere categorieën kunnen worden onderverdeeld, vb. *een hoedje* of *een slag in de rondte*. In de synchrone data hebben we het gebruik van de constructie vergeleken in Belgisch versus Nederlands Nederlands. In het algemeen zijn de gelijkenissen tussen de twee nationale variëteiten van het Nederlands groter dan de verschillen. Heel wat werkwoorden en intensiveerders, alsook specifieke werkwoord-intensiveerdercombinaties, komen zowel voor in Belgisch als Nederlands Nederlands, al lijken

beide variëteiten er wel enkele eigen voorkeuren op na te houden. Zo worden de intensiveerders *rot*, *suf* en *kapot* veel frequenter gebruikt in het Nederlands Nederlands en komen *te pletter* en *(zich) uit de naad (werken)* opvallend vaker voor in het Belgisch Nederlands. Los daarvan zijn er ook enkele nationaal-exclusieve intensiveerders of uitdrukkingen, d.w.z. dat ze vrij vaak voorkomen in de ene variëteit maar geheel afwezig zijn in de andere. Kenmerkend voor Nederlands Nederlands zijn bijvoorbeeld de intensiveerder *een slag in de rondte* en de uitdrukkingen *zich wild schrikken/ergeren*; in het Belgisch Nederlands vinden we o.a. de exclusieve intensiveerders *de ziel uit het lijf* en *zot* en de uitdrukking *zich steendood vervelen*.

Als we de diachrone data van Delpher bekijken, wordt meteen duidelijk dat de constructie heel wat veranderingen heeft ondergaan sinds het begin van de 19de eeuw. De constructie is niet alleen veel frequenter geworden (genormaliseerde frequentie van 1.17 per tien miljoen woorden in de jaren 1830 versus 68.51 per tien miljoen woorden in de jaren 1990), ze wordt ook gebruikt met een steeds grotere variatie aan werkwoorden en intensiveerders (4 INT, 5 V in de jaren 1830 versus 115 INT, 131 V in de jaren 1990). Als we de algemene frequentieontwikkeling volgen, blijkt zowel de token- als type-expansie rond het midden van de 20ste eeuw een versnelling hoger te schakelen. Uit de individuele ontwikkeling van enkele werkwoorden en intensiveerders kunnen we afleiden dat de algemene toename in tokenfrequentie vooral wordt getrokken door een aantal hoogfrequente werkwoorden en intensiveerders. Bij de werkwoorden gaat het om enkele emotieve en fysieke handelingswerkwoorden die al sinds de 19de eeuw prominent figureren in de constructie (*lachen, zich ergeren, zich vervelen, zich schamen, werken, lopen* en sinds 1910 ook *schrikken*). Hoewel het werkwoordslot wel een semantische expansie heeft ondergaan (allerlei activiteiten kwamen bijvoorbeeld nog niet voor in de 19de en vroege 20ste eeuw), lijkt er toch enige continuïteit te zijn met betrekking tot de werkwoorden waar de constructie een voorkeur voor vertoont. Bij de intensiveerders zijn er duidelijk meer historische verschuivingen. Slechts een aantal van de intensiveerders die momenteel erg frequent zijn, kwamen al voor in de 19de eeuw (*dood, suf* en *het vuur uit de sloffen*); de andere zijn duidelijk recentere succesverhalen die pas in de tweede helft van de 20ste eeuw zijn opgedoken in de constructie (vb. *rot, uit de naad, te pletter*). In het algemeen merken we ook bij het intensiveerderslot een duidelijke semantische uitbreiding: niettegenstaande de occasionele intensiveerder van "onvervreemdbaar bezit", drukten de oudste intensiveerders bijna uitsluitend een negatieve toestand uit. De eerste kleurtermen duiken pas op aan het begin van de 20ste eeuw en de categorie van ziektetermen die in het hedendaags Nederlands zo gevarieerd is, begint pas vanaf het midden van de 20ste eeuw heel wat nieuwe leden aan te trekken. Kortom, het lijkt erop alsof de constructie op alle vlakken een enorme expansie heeft ondergaan in de afgelopen twee eeuwen, al zal blijken dat dat beeld enigszins genuanceerd moet worden.

2. Collocationele patronen. Hoewel we de intensiveerders en de werkwoorden tot nu toe als twee onafhankelijke slots in de constructie hebben beschouwd, blijken er belangrijke interacties te zijn tussen beide slots. Het is namelijk zo dat niet alle werkwoorden en intensiveerders zich even flexibel opstellen in de constructie; sommige vertonen erg specifieke voorkeuren met betrekking tot de items waarmee ze gecombineerd worden. Er lijkt in dat geval sprake te zijn van coselectie, waarbij de keuze voor een bepaald item een reeks mogelijkheden (en "onmogelijkheden") op het andere slot projecteert. Natuurlijk is het niet zo dat elke werkwoord-intensiveerdercombinatie die in ons corpus niet geattesteerd is, per definitie ook onmogelijk is: er zijn heel wat combinaties

die niet voorkomen maar op zich niet meteen vreemd klinken. Anderzijds zijn er ook combinaties die wél geattesteerd zijn maar toch duidelijk onconventioneel klinken. Combinaties als *zich een hoedje tappen* of *zich blauw klagen* springen meteen in het oog omdat de taalgebruiker zo vertrouwd is met de vaste combinaties *zich een hoedje schrikken* en *zich blauw betalen/ergeren* dat hij/zij die specifieke intensiveerders niet meteen met een ander werkwoord zou verwachten. Gezien de expressieve kracht van de constructie is het natuurlijk mogelijk dat taalgebruikers bewust de grenzen van bepaalde conventies opzoeken of bepaalde restricties gaan overtreden om extra effect te creëren.

We bekijken ook hoe het collocationele gedrag van bepaalde werkwoorden en intensiveerders zich doorheen de tijd heeft ontwikkeld. Zo merken we dat bepaalde werkwoorden of intensiveerders in de constructie geïntroduceerd werden als deel van een vaste uitdrukking, maar na verloop van tijd ook met andere items gecombineerd konden worden (vb. *zich het vuur uit de sloffen lopen*). Er zijn echter ook enkele werkwoorden en intensiveerders die al meerdere decennia min of meer beperkt zijn tot één of twee vaste uitdrukkingen (vb. *zich* ***een hoedje*** *schrikken, zich* ***groen en geel*** *ergeren, zich suf* ***piekeren***). In uitzonderlijke gevallen kan het zelfs gebeuren dat items die vroeger een grotere combinatorische flexibiliteit of wijdere toepasbaarheid genoten, zich gaan terugtrekken tot enkele conventionele collocaties (vb. *wild* in *zich wild ergeren/schrikken* of *blauw* in *zich blauw betalen/ergeren*). Dergelijke veranderingen in het collocationele gedrag van de individuele items heeft implicaties voor de productiviteit van die items, die hieronder besproken wordt.

3. *Productiviteit.* De productiviteit van een constructie verwijst naar de "extensibiliteit" van de constructie, d.w.z. de mate waarin het mogelijk is om ze uit te breiden naar nieuwe types. Hoewel de constructie [SUBJ V REFL INT] in het algemeen zeer productief is, blijkt uit de vorige paragraaf dat er toch wel wat verschillen zijn op het niveau van de individuele werkwoorden en intensiveerders. Als we aannemen dat ieder werkwoord en iedere intensiveerder op een lager niveau ook een subconstructie vormt, van het type [SUBJ *specifiek werkwoord* REFL INT] en [SUBJ V REFL *specifieke intensiveerder*], kunnen we ook op dat niveau de productiviteit meten van de open werkwoord- en intensiveerderslots. We doen dat aan de hand van een aantal specifieke productiviteitsmaten, die in het algemeen gebaseerd zijn op het idee dat de productiviteit van een constructie positief beïnvloed wordt door een hoge typefrequentie en een hoge proportie aan hapaxen (Baayen 1990, 1992, 1993, 2009, Baayen & Lieber 1991). Daarnaast besteden we ook aandacht aan de semantische aspecten van productiviteit. Volgens Barðdal (2008) is het namelijk mogelijk dat ook een constructie met een lagere typefrequentie een zekere productiviteit vertoont, tenminste als de types een hoge graad aan semantische coherentie vertonen. Dit idee wordt weergegeven in een productiviteitscontinuüm dat aan het ene uiteinde wordt gekenmerkt door hoge typefrequentie en lage coherentie en aan het andere uiteinde door lage typefrequentie en hoge coherentie. Zelfs als er slechts één type voorkomt kan de constructie binnen dit model via analogische extensies toch uitgebreid worden naar nieuwe types (als dat ene type heel frequent is). De frequentiematen en semantische aspecten worden samengenomen in een multidimensioneel productiviteitsmodel. In het hedendaags Nederlands vinden we een aantal werkwoorden (*lachen*, *lopen*, *schrikken*, *werken*, *zich ergeren*...) en intensiveerders (*kapot, rot, suf, te pletter, wezenloos*...) die vrij productief zijn, in die zin dat ze voorkomen met een groot aantal types uit verschillende semantische klassen. De werkwoorden en intensiveerders die bijna uitsluitend

voorkomen in vaste collocaties, daarentegen, kunnen niet of nauwelijks productief worden genoemd (vb. de intensiveerders *een hoedje*, *groen en geel*, *wild* of het werkwoord *piekeren*). Tussen deze twee categorieën in zijn er een aantal items die enige productiviteit vertonen binnen een beperkt semantisch domein. De intensiveerders *de longen uit het lijf, de ziel uit het lijf* en *uit de naad*, bijvoorbeeld, komen uitsluitend voor met werkwoorden die ofwel een fysieke handeling (*lopen, fietsen…*) ofwel de productie van een (luid) geluid (*zingen, schreeuwen…*) uitdrukken. Tegelijk zijn er echter enkele intensiveerders die zich niet perfect laten inpassen in het productiviteitsmodel van Barðdal (2008). De voorspelling dat constructies met een lage typefrequentie enkel productief kunnen zijn als hun types een hoge graad aan semantische coherentie vertonen lijkt niet altijd te worden bevestigd door onze data, die ook verschillende productiviteitseilanden bevatten die weinig semantische coherentie vertonen.

Vanuit diachroon perspectief is de productiviteit van de subconstructies onderhevig aan veranderingen. De intensiveerder die de grootste ontwikkeling heeft doorgemaakt is zonder twijfel *suf*: tot het midden van de 20ste eeuw kwam de intensiveerder enkel voor met werkwoorden die een mentale activiteit uitdrukken, maar in hedendaags Nederlands Nederlands heeft *suf* zich ontpopt tot de meest flexibele intensiveerder bij uitstek, met maar liefst 61 verschillende werkwoordstypes uit allerlei semantische categorieën. Daarnaast is het mogelijk dat de productiviteit van specifieke subconstructies afneemt, ook wanneer de productiviteit van de constructie op het allerhoogste niveau lijkt toe te nemen. In sommige gevallen verdwijnt de intensiveerder volledig (vb. *een aap* in hedendaags Nederlands Nederlands), in andere blijft die nog bewaard in enkele vaste collocaties – als zogezegde "overblijfselen" van de vroegere productiviteit. Een duidelijk voorbeeld daarvan is de intensiveerder *wild*. Tussen de jaren 1950 en 1980 werd de intensiveerder gebruikt met een kleine set van werkwoorden uit verschillende semantische klassen. Geleidelijk aan zijn er twee conventionele collocaties ontstaan, *zich wild ergeren* en *zich wild schrikken*, die er uiteindelijk in geslaagd zijn om alle andere werkwoorden te verdringen. De gradaties en verschuivingen in productiviteit die hier werden besproken, spelen een belangrijke rol in de (re)organisatie van het constructionele netwerk, zie hieronder.

4. Het constructionele netwerk. Het feit dat de constructie een zekere mix tussen productiviteit en conventionaliteit vertoont, kan worden weergegeven in de vorm van een constructioneel netwerk. Binnen die hiërarchie is het mogelijk om bepaalde idiosyncratische restricties te poneren op lagere niveaus die niet van toepassing zijn op de constructie in het algemeen. Dat netwerk wordt bottom-up opgebouwd vanuit lexicaal-specifieke, concrete instanties op het laagste niveau naar steeds abstractere patronen op hogere niveaus. Concreet wordt er pas een abstracter patroon aangenomen wanneer daar voldoende bewijs voor bestaat in de data. Als een bepaalde intensiveerder of een bepaald werkwoord uitsluitend voorkomt met één of twee andere items, kunnen we voor dat specifieke item geen tussenliggend subschema aannemen. Bijvoorbeeld, het netwerk bevat hoogstwaarschijnlijk geen subschema [SUBJ V REFL *groen en geel*] aangezien *groen en geel* enkel voorkomt met *zich ergeren*. Die conventionele collocatie zit "vast" op het micro-constructieniveau als [SUBJ *ergeren* REFL *groen en geel*]. Voor een intensiveerder als *suf* daarentegen kunnen we wel stellen dat er een abstract patroon [SUBJ V REFL *suf*] bestaat, aangezien *suf* met een waaier aan verschillende werkwoorden voorkomt. Het subschema abstraheert m.a.w. over het specifieke werkwoord. Ook voor de intensiveerder *uit de naad* kunnen we zo'n subschema aannemen, maar aangezien de werkwoorden allemaal tot een beperkt

semantisch domein behoren, moeten we wel nog een semantische restrictie toevoegen aan het werkwoordslot, vb. [SUBJ V$_{\text{fysieke handeling/productie luid geluid}}$ REFL *uit de naad*]. Dat subschema zit dus op een lager niveau in het netwerk dan [SUBJ V REFL *suf*] omdat het meer gespecificeerd is. Op die manier ontstaat er een complexe structuur van micro-constructies en subschemas op verschillende niveaus van abstractie. De huidige structuur van het netwerk is het resultaat van de diachrone ontwikkelingen die de constructie en haar subconstructies hebben ondergaan. Wanneer één of beide elementen van een conventionele collocatie hun collocationele reikwijdte gaan uitbreiden naar nieuwe elementen, ontstaat er een productief subschema. In de vroege stadia van productiviteit zullen de nieuwe types nog sterk semantisch verwant zijn aan de oorspronkelijke collocaten (Suttle & Goldberg 2011, Zeschel 2012). Zodra het subschema gevormd is, kan het nog meer types uit andere semantische domeinen gaan aantrekken, wat ervoor zorgt dat het subschema opschuift naar een hoger niveau in het netwerk. We moeten het netwerk zien als een dynamisch geheel dat bij ieder gebruik van de constructie licht gewijzigd wordt. Na verloop van tijd kunnen al die kleine wijzigingen tot grote reorganisaties van het netwerk leiden.

Afhankelijk van het soort generalisaties of abstracties die er gemaakt worden, kunnen we verschillende mogelijke representaties van het netwerk bouwen. Als we de focus bij het werkwoord in plaats van de intensiveerder leggen, krijgen we bijvoorbeeld subschema's zoals [SUBJ *ergeren* REFL INT]. Op die manier kunnen bepaalde micro-constructies die in één representatie geïsoleerd lijken te zijn, in een andere representatie vaak wel gemotiveerd worden door een tussenliggend subschema (vb. de exclusieve associatie van *groen en geel* met *zich ergeren* geeft geen aanleiding tot [SUBJ V REFL *groen en geel*] maar wordt wel gemotiveerd door een subschema [SUBJ *ergeren* REFL INT]). Een dergelijke multiconfigurationele of multirepresentationele benadering is belangrijk als we aan de constructionele netwerken, als "theoretische abstracties", ook een zekere cognitieve realiteit willen toekennen. De taalgebruiker is er namelijk toe in staat om verschillende soorten generalisaties te maken, wat erop neerkomt dat al onze mogelijke representaties van het netwerk voor de taalgebruiker eigenlijk samenvloeien in één dynamisch, interactief geheel.

Conclusie

Dit onderzoek startte vanuit de observatie dat de intensiverende pseudoreflexieve resultatiefconstructie in het hedendaags Nederlands een interessante mix van productiviteit en conventionaliteit vertoont. We hebben aangetoond dat deze synchrone variatie kan worden weergegeven aan de hand van een hiërarchisch georganiseerd constructioneel netwerk. Binnen dit netwerk is er plaats voor zowel productiviteit op het hoogste niveau, conventionele collocaties op het laagste niveau en verschillende graden van productiviteit op meerdere subschemaniveaus tussenin. Uit de analyse van de diachrone data blijkt dat verschillende belangwekkende constructionele veranderingen niet altijd zichtbaar zijn op het meest schematische niveau van de constructie. In het geval van de intensiverende pseudoreflexieve resultatiefconstructie, [SUBJ V REFL INT], blijken de meest interessante verschuivingen zich vooral voor te doen op het niveau van de individuele werkwoorden en intensiveerders (d.w.z. de gedeeltelijk gespecifieerde subschema's) en op het niveau van specifieke werkwoord-intensiveerdercombinaties (d.w.z. de lexicaal gespecificeerde micro-constructies). Aan de hand

van het constructionele netwerk is het mogelijk om allerlei verschuivingen op verschillende niveaus van abstractie te volgen.

Aangezien er een nauwe band wordt aangenomen tussen schematiciteit en productiviteit, is er voor dat laatste een belangrijke rol weggelegd bij de (re)organisatie van het netwerk. We hebben gezien dat werkwoorden of intensiveerders die slechts een zeer beperkte collocationele reikwijdte hebben, niet als productief kunnen worden beschouwd en dat ze binnen het netwerk vastzitten op het micro-constructieniveau. Ook bij de subschema's hangt de positie samen met de productiviteit van het schema, in die zin dat een subschema dat met veel types voorkomt en onderhevig is aan weinig restricties schematischer is (d.w.z. op een hoger niveau in het netwerk zit) dan een subschema dat slechts een beperkte toepasbaarheid heeft. Deze studie heeft ook aangetoond dat productiviteit zelf een complex, multi-gefaceteerd fenomeen is waarbij verschillende kwantitatieve en kwalitatieve factoren een rol spelen. Aan de hand van een multidimensioneel productiviteitsmodel was het mogelijk om een vrij nauwkeurig beeld te schetsen van hoe productiviteit op verschillende niveaus van abstractie aan het werk is, maar tegelijk bleek dat het model nog op verschillende vlakken verfijnd kan worden.

Verder heeft dit onderzoek aangetoond dat de expressieve betekeniscomponent van de constructie een zekere invloed heeft op haar gebruik en diachrone ontwikkeling. We kunnen stellen dat de drang naar expressiviteit vermoedelijk heeft bijgedragen tot de creativiteit en de productiviteit van de constructie op het meest abstracte niveau: taalgebruikers komen geregeld met nieuwe intensiveerders (of variaties op bestaande intensiveerders) op de proppen. Op die manier heeft het repertoire van intensiveerders zich gestaag uitgebreid en zijn er in het hedendaags Nederlands meer dan 120 verschillende intensiveerders "in omloop". Op een lager niveau zijn er echter ook sporen van interne rivaliteit tussen de intensiveerders. Wanneer een bepaalde intensiveerder te frequent is geworden, kan die een deel van zijn expressieve kracht verliezen en in sommige situaties vervangen worden door een recenter alternatief. In veel gevallen behoudt de oudere intensiveerder nog wel een zekere productiviteit, maar soms trekt die zich terug tot enkele vaste collocaties – en in het extreme geval kan de intensiveerder zelfs geheel verdwijnen. Kortom, de expressiviteit van de constructie zorgt voor een zekere gelaagdheid van zowel oude, vaste waarden als nieuwe creatieve vondsten. Dat proces van innovatie en hernieuwing leidt bovendien tot een constante reorganisatie van de structuur van het netwerk.

Summary

Theoretical framework and methodology

This thesis presents the results of a synchronic and diachronic corpus investigation into the role of productivity in (re)shaping the constructional network. We first examine the factors that come into play when determining productivity and the tools that are at our disposal to measure the impact of those factors. It is then investigated to what extent productivity (at different levels of abstraction, including schema-level productivity and low-level analogy) has played a role in the organisation and internal shifts of the constructional network. In order to do so, we focus on one specific construction, viz. the intensifying fake reflexive resultative construction, as illustrated in the following examples:

(287) Overtreders, mensen die beboetbaar bezig zijn geweest, schrikken zich vaak een hoedje over de hoogte van de boete. (SoNaR)
[...] startle themselves often a little hat [...]
'Offenders, people who have committed a punishable offence, are often highly startled by the amount of the fine.'

(288) Ik zapte er nu langs en schrok me de blaren van het geplamuurde gezicht met die rode lippen (Twitter, 23/10/2016)
[...] startled myself the blisters off the plastered face [...]
'I swooped by and was highly startled by those red lips.'

In present-day Dutch, the construction displays an interesting mix of productivity and lexical idiosyncrasy: even though the construction seems to allow for a lot of linguistic creativity, see (288), there are also a number of "fixed expressions" or collocational preferences that keep the creativity within bounds, see (287). In this thesis, we trace the recent history of the construction in order to find out which constructional changes it has undergone since the early 19th Century. Given the intensifying meaning component, the construction also allows us to explore the potential influence of expressivity on the observed variation and changes. Several studies have shown that language users, especially within the linguistic domain of intensification, appear to have a certain universal need for expressivity, in the sense that they (consciously or unconsciously) wish to set themselves apart from other language users by means of their language use. In present-day Dutch, this drive for expressivity appears to feed into linguistic creativity, but it may also be of importance from a diachronic point of view. When a given linguistic expression becomes too frequent or too widespread in the linguistic community, it may

shed some of its expressive force. In some contexts, the expression may then be replaced by – or at least complemented with – a new, more expressive alternative. This pragmatic wear-and-tear can lead to innovation and renewal, as new forms are constantly being introduced (Stoffel 1901, Bolinger 1972, Partington 1993, Lorenz 2002, De Clerck & Colleman 2013).

Concretely, we searched a large journalistic corpus, compiled on the basis of the historical newspaper data from Delpher (for the period 1830-1995, Netherlandic Dutch only) and the present-day newspapers in SoNaR (from 1995 onwards, Belgian and Netherlandic Dutch) in order to retrieve an exhaustive set of examples of the intensifying fake reflexive resultative construction. We used a cyclic search procedure, in which the output of one search query was used as input for the next search query. After multiple rounds and extensive manual filtering, we composed a data set of 3,487 examples for present-day Dutch (1,042 for Netherlandic Dutch, 2,445 for Belgian Dutch) and 6,137 examples for the period from 1830 to 1995. The entire data set was annotated for several linguistic variables, i.e. lemma verb, reflexivity and transitivity of the verb, lemma intensifier, syntactic category of the intensifier and form of the reflexive pronoun, as well as one extra-linguistic variable, viz. national variety.

Results

1. *General frequency, use and development.* A first look at the synchronic data tells us that the construction shows a high degree of variability in present-day Dutch, with a total of 260 verbs and 122 different intensifiers in the entire SoNaR data set. The fact that about half of the verb types and over a third of the intensifier types occur only once (i.e. the so-called hapax legomena), testifies to the creative potential of the construction. At the same time, we find that specific individual verbs and intensifiers (and verb-intensifier combinations, cf. infra) occur with a much higher frequency than others, which shows that there is also a large degree of conventionality involved. If we look at the verb semantics, it appears that the verbs can denote all kinds of activities which have an inherent aspect that can be intensified in one way or another – although there appears to be a certain preference for experience verbs and physical activity verbs. The intensifiers can take different syntactic forms, e.g. AP, NP, PP, NP+PP or NP+AP and also display a certain semantic variability. A lot of highly frequent intensifiers express a negatively connoted state (e.g. *dood* 'dead', *suf* 'drowsy', *te pletter* 'to smithereens', *uit de naad* 'out of the seam'). In addition, there are a number of intensifiers that involve a body-part or piece of clothing (e.g. *de longen uit het lijf* 'the lungs out of the body', *de ziel uit het lijf* 'the soul out of the body', *het vuur uit de sloffen* 'the fire out of the slippers'), an extraordinarily variegated group of disease terms (e.g. *de pleuris* 'the pleurisy', *de tering* 'the consumption", *de tyfus* 'the typhoid', *het apelazerus* 'fictitious disease', *het leplazerus* 'fictitious disease', *het schompes* 'fictitious disease') and a couple of colour terms (e.g. *blauw* 'blue', *groen en geel* 'green and yellow', *groen* 'green'). With the exception of the last category, it appears that most intensifiers are recruited from conceptual domains that have some kind of negative connotation. Finally, there are a number of "isolated" intensifiers that cannot easily be categorised into one of the previously established groups, e.g. *een hoedje* 'a little hat' or *een slag in de rondte* 'a punch around'. In the synchronic data, we also compared the use of the construction between Belgian and Netherlandic Dutch. On the whole, the similarities between both national varieties outweigh the differences. A lot of verbs and intensifiers, as well as specific verb-intensifier combinations, are used in both Belgian and Netherlandic Dutch, although

speakers of both varieties seem to display slightly different preferences. For example, the intensifiers *rot* 'rotten', *suf* 'drowsy' and *kapot* 'broken' are used significantly more often in Netherlandic Dutch, whereas *te pletter* 'to smithereens' and *(zich) uit de naad (werken)* '(to work oneself) out of the seam' are much more frequent in Belgian Dutch. There are also a couple of truly nationally-exclusive intensifiers and idiomatic expressions, which are quite frequent in one of the two national varieties but wholly absent in the other. Typical for Netherlandic Dutch are, for example, the intensifier *een slag in de rondte* 'a punch around' and the expressions *zich wild schrikken/ergeren* 'to startle/annoy oneself wild'; in Belgian Dutch, then, we find the exclusive intensifier *de ziel uit het lijf* 'the soul out of the body' and the expression *zich steendood vervelen* 'to bore oneself stone-dead'.

If we take a look at the diachronic Delpher data, we immediately find that the construction has undergone many changes since the early 19th Century. Not only has the construction increased its frequency of use (normalised frequency of 1.17 per ten million words in the 1830s versus 68.51 per ten million words in the 1990s), it has also expanded and diversified the range of verbs and intensifiers that can be used in its slots (4 INT, 5 V in the 1830s versus 115 INT, 131 V in the 1990s). A closer look at the general frequency development indicates that both the token expansion and the type increase have taken it up a notch around the 1930s. Tracking the development of some individual verbs and intensifiers, we observe that the general expansion is carried by a number of highly frequent verbs and intensifiers. In the verb slot, this mainly concerns a number of experience verbs and physical activity verbs that have been prominent in the construction since the 19th Century (*lachen* 'to laugh', *zich ergeren* 'to be annoyed', *zich vervelen* 'to be bored', *zich schamen* 'to be embarrassed', *werken* 'to work', *lopen* 'to run' and since the 1910s also *schrikken* 'to be startled'). Even though the verb slot does display signs of semantic expansion (in the sense that a lot of activities were not yet attested in the 19th and early 20th Century), it does show remarkable diachronic continuity. In the intensifier slot, there appear to be more substantial diachronic shifts. Only a fraction of the intensifiers that are currently highly frequent were already used in the 19th Century (*dood* 'dead', *suf* 'drowsy' and *het vuur uit de sloffen* 'the fire out of the slippers'); the others are recent success stories that only joined the construction in the second half of the 20th Century (e.g. *rot* 'rotten', *uit de naad* 'out of the seam', *te pletter* 'to smithereens'). In general, the intensifier slot has also widened its semantic scope: with the exception of the occasional inalienable possession intensifier, most of the oldest intensifiers denote a negatively connoted state. The first colour terms are only introduced in the early 20th Century and the category of disease terms that is so prolific in present-day Dutch only starts to build its extensive repertoire around the mid-20th Century. In sum, it seems as if the construction has undergone an enormous expansion on all fronts over the past 200 years or so – although it will be shown below that we need to add some nuance to this image of "general expansion".

2. Collocational patterns. Up to this point, we have treated the verbs and intensifiers as two independent slots in the construction. However, there are important interactions between both slots: not all verbs and intensifiers show the same degree of combinatorial flexibility and some even display very specific preferences with respect to the items they are paired up with. We seem to be dealing with some kind of coselection, in which the choice for one particular item projects a range of possibilities (and "impossibilities") onto the other slot. Of course, it is not the case that a verb-intensifier combination that happens not to be attested in our corpus is by definition also

an impossible combination: there are several unattested (but possible) verb-intensifier combinations that do not sound odd at all. At the same time, there are a number of verb-intensifier combinations that *are* attested and that do sound rather unconventional. Collocations like *zich een hoedje tappen* 'to tap oneself a little hat' and *zich blauw klagen* 'to complain oneself blue' immediately jump to the eye because the native speaker of Dutch is so familiar with the fixed expressions *zich een hoedje schrikken* 'to startle oneself a little hat' and *zich blauw betalen/ergeren* 'to pay/annoy oneself blue' that he/she may not expect these intensifiers to be used with any other verbs at all. Given the expressive nature of the construction, of course, it is possible that users deliberately flirt with the edges of conventionality or even override certain restrictions in order to create an extra effect.

We also want to see how the collocational patterns of certain verbs and intensifiers have developed over time. We find that several verbs and intensifiers were introduced in the construction as part of a fixed expression, but gradually also started to be used outside of that particular collocation (e.g. *zich het vuur uit de sloffen lopen* 'to run oneself the fire out of the slippers'). At the same time, there are a number of verbs and intensifiers that have been virtually limited to one or two fixed collocations for several decades (e.g. *zich* ***een hoedje*** *schrikken* 'to startle oneself a little hat', *zich* ***groen en geel*** *ergeren* 'to annoy oneself green and yellow' or *zich suf* ***piekeren*** 'to annoy oneself drowsy'). Exceptionally, we even find items that used to have a higher degree of combinatorial flexibility retreating to particular collocations (e.g. *wild* 'wild' in *zich wild ergeren/schrikken* 'to annoy/startle oneself wild' or *blauw* 'blue' in *zich blauw betalen/ergeren* 'to pay/annoy oneself blue'). Such changes in the collocational behaviour of individual items have certain implications for the productivity of those items, as will be discussed below.

3. *Productivity.* The productivity of a construction is defined as the "extensibility" of the construction, i.e. the extent to which it is possible to extend it to new types. Even though the construction at the maximum level of schematicity [SUBJ V REFL INT] is very productive in general, the previous paragraph suggests that we get a slightly different image if we consider the level of individual verbs and intensifiers. Each verb and intensifier can be assumed to form a subconstruction at a lower level, i.e. [SUBJ *specific verb* REFL INT] or [SUBJ V REFL *specific intensifier*]. In order to measure the productivity at different levels of abstraction, we are using a number of productivity measures, which are based on the idea that productivity is positively influenced by a high type frequency and a high number of hapaxes (Baayen 1990, 1992, 1993, 2009, Baayen & Lieber 1991). In addition, we address the impact of semantics on productivity. Barðdal (2008) argues that it is possible for a construction that has a lower type frequency to still display a certain degree of productivity, provided that the types show a high degree of semantic coherence. This idea translates to a productivity continuum, characterised by high type frequency and low coherence on one end and low type frequency and high coherence on the other. Even if a pattern is instantiated by only one type, it can still be extended to new types through analogical extensions (if the one type is highly token frequent). The frequency-based measures and semantic aspects are taken together in a multidimensional productivity model. In present-day Dutch, we find a number of verbs (*lachen* 'to laugh', *lopen* 'to run', *schrikken* 'to be startled', *werken* 'to work', *zich ergeren* 'to be annoyed'...) and intensifiers (*kapot* 'broken', *rot* 'rotten', *suf* 'drowsy', *te pletter* 'to smithereens', *wezenloos* 'vacant'...) that are quite productive, in the sense that they can occur with a wide array of types from different semantic classes. The

verbs and intensifiers that were found to be near-exclusively used in a number of fixed expressions, however, cannot (or barely) be said to be productive at all (e.g. the intensifiers *een hoedje* 'a little hat', *groen en geel* 'green and yellow', *wild* 'wild' or the verb *piekeren* 'to worry'). In between those two categories there are a number of items that show some signs of productivity within a delimited semantic domain. The intensifiers *de longen uit het lijf* 'the lungs out of the body', *de ziel uit het lijf* 'the soul out of the body' and *uit de naad* 'out of the seam', for example, are limited to verbs that either express a physical activity (running, cycling...) or the production of a loud noise (singing, screaming...). At the same time, we also find a number of specific intensifiers that do not straightforwardly fit onto the cline of Barðdal (2008). The prediction that (sub)constructions with a low type frequency can only be productive if their types show a high degree of semantic coherence does not always appear to be borne out by our data, as we also find several productivity islands that show very little internal coherence.

From a diachronic point of view, the degree of productivity of the subconstructions is subject to change. The intensifier that has without doubt undergone the most drastic expansion is *suf* 'drowsy': until the mid-20th Century, the intensifier was exclusively used with verbs that denote a mental activity, but in present-day Netherlandic Dutch, it has developed into the most productive intensifier by a large margin, occurring with 61 verb types from different semantic classes. We also observed that, even when the productivity of the construction at the maximum level of schematicity appears to be increasing, subschemas at lower levels may decrease in productivity or even cease to be productive. In some cases, the item disappears from the construction entirely (e.g. *een aap* 'a monkey' in present-day Netherlandic Dutch), but in other cases the item survives in a number of conventional collocations – which serve as relics of its former productivity. A clear example of the latter scenario is *wild* 'wild'. Between the 1950s and 1980s, the intensifier was used with a (small) set of verbs from a number of different semantic classes. Gradually, two conventional collocations started arising, viz *zich wild ergeren* 'to annoy oneself wild' and *zich wild schrikken* 'to startle oneself wild', which managed to oust all other verbs. The variation and shifts in productivity that were just discussed play an important role in the (re)organisation of the constructional network, see below.

4. The constructional network. The fact that the construction displays a certain mix of productivity and conventionality is captured in a constructional network. Within that hierarchy, it is possible to posit certain restrictions at lower levels that do not operate on the construction as a whole. The network is built bottom-up, starting with lexically-specific concrete instances at the bottom, and further abstracting upwards to increasingly more schematic levels if the data support the existence of such a higher-level generalisation. If a certain verb or intensifier is exclusively used with one or two other items, we cannot assume the existence of an intermediate subschema. For instance, the network most likely does not contain a subschema [SUBJ V REFL *groen en geel*], given that *groen en geel* 'green and yellow' is exclusively combined with the verb *zich ergeren* 'to be annoyed'. In other words, the conventional collocation is "stuck" at the micro-construction level as [SUBJ *ergeren* REFL *groen en geel*]. For an intensifier such as *suf* 'drowsy', on the other hand, which is found to combine with an array of semantically diverse verb types, we can posit a more abstract pattern [SUBJ V REFL *suf*] that abstracts away from the specific verb. For the intensifier *uit de naad* 'out of the seam', as well, we could posit an intermediate subschema, but given that all of its verbal collocates belong to a delimited semantic domain, we need to add

some kind of semantic restriction on the verb slot, e.g. [SUBJ $V_{\text{physical effort/noise emission}}$ REFL *uit de naad*]. This subschema is positioned at a lower level in the hierarchy than [SUBJ V REFL *suf*] because it is more specific (i.e. less schematic). This process of schema-formation results in an intricate structure of micro-constructions and subschemas at different levels of abstraction. The current structure of the network is the result of the diachronic changes the construction and its subconstructions have undergone. When one or both elements of a conventional verb-intensifier combination start expanding their collocational range to new types, a (partially) productive subschema may arise. In the early stages of productivity, these new coinages will be highly semantically related to the original collocate (Suttle & Goldberg 2011, Zeschel 2012). As soon as the subschema has been formed, it can come to attract even more types from other semantic domains, causing the subschema to increase its schematicity and shift upwards to a higher level in the network. We should think of the network as a dynamic system that is slightly modified upon each use of the construction. Over time, these kinds of little shifts can lead to substantial reorganisations within the network structure.

We further suggest that it is possible to build multiple representations of one constructional network, depending on the kinds of generalisations or abstractions that are made. If the network is centred on the verb slot instead of the intensifier slot, we get subschemas like [SUBJ *ergeren* REFL INT]. A micro-construction that appears to be isolated in one representation of the network, can then be perfectly motivated by a low-level subschema within another possible representation (e.g. the exclusive association of *groen en geel* 'green and yellow' and *zich ergeren* 'to be annoyed' may not give rise to a subschema [SUBJ V REFL *groen en geel*], but it is of course motivated by the subschema [SUBJ *ergeren* REFL INT]). Such a multiconfigurational or multirepresentational approach is important if we want to add some cognitive reality to the constructional networks, rather than just viewing them as theoretical constructs. As the language user is able to make different kinds of generalisations at the same time, all our possible representations of the network come together in one dynamic, interactive system.

Conclusion

This investigation started out from the observation that the intensifying fake reflexive resultative construction in present-day Dutch displays an interesting mix of productivity and conventionality. We have shown that this synchronic variation can be represented in a taxonomically organised constructional network; a Lexicality-Schematicity Hierarchy, if you will. This network straightforwardly accommodates both productivity at the highest level of abstraction, "fixed expressions" at the lowest level and varying degrees of productivity at multiple intermediate levels. Based on the analysis of the diachronic data, we argue that there are several important constructional changes that are not visible at the most schematic level of the construction. In the case of the intensifying fake reflexive resultative construction [SUBJ V REFL INT], the more interesting changes appear to be taking place at the level of specific verbs and intensifiers (i.e. the level of partially specified subschemas) and at the level of specific verb-intensifier combinations (i.e. the level of lexically specified micro-constructions). The constructional network allows us to track the different kinds of shifts at multiple levels of abstraction.

Given the assumed tight interrelatedness between schematicity and productivity, it makes sense that productivity plays an important role in the (re)organisation of the network. We have shown that verbs or intensifiers with a limited collocational range cannot be said to be productive and, accordingly, they are only represented at the micro-construction level in the network. With respect to subschemas, as well, their position in the hierarchy is linked to their degree of productivity, in the sense that a subschema that may host a large variety of types and is not subject to any obvious restrictions is more schematic (i.e. situated at a higher level in the network) than a subschema that only has a limited range of application. This study has also illustrated that productivity in itself is a complex, multifaceted phenomenon that may be influenced by different quantitative and qualitative factors. On the basis of a multidimensional productivity model, we were able to sketch a rather detailed picture of how productivity is important at different levels of granularity and abstraction, but we also concluded that the model can be refined in a number of ways.

Finally, this investigation has shown that the expressive meaning component of the construction exerts some influence on its use and development. We can propose that the need for expressivity has fuelled the creative and productive use of the schematic construction: language users are constantly inventing and introducing new intensifiers (or new variants of existing intensifiers). In doing so, they have gradually expanded the repertoire of intensifiers, to the point where we find over 120 different intensifiers being used in present-day Dutch. At a lower level, we also find indications of a power struggle or competition between the intensifiers. If one intensifier has become too frequent, it may lose some of its expressive force and be replaced by a more recent alternative. In many cases, the older intensifier still retains some degree of productivity, but sometimes it retreats to specific collocations – or, in the most extreme scenario, it drops out of use entirely. Summing up, the expressivity of the construction translates to a certain layering of both old and new, creative intensifiers; this process of innovation and renewal further contributes to the constant reshaping of the constructional network.